Speech Recognition System (Home Appliances Controller of Local & Remote System) using LPC & HMMs Methodologies

Nadeem Ahmed Kanasro[1], Shazia Samoon[2], Mujeeb-U-Rehman Jamali[3], Fiaz Ahmed Memon[4],
Shahmurad Chandio[5]
[1]IMCS & SUCL, University of Sindh, Jamshoro, Pakistan, nadeem.kanasro@usindh.edu.pk
[2]SUCL, University of Sindh, Jamshoro, Pakistan, shazia.samoon@usindh.edu.pk
[3]IMCS & SUCL, University of Sindh, Jamshoro, Pakistan, mujeebjamali@usindh.edu.pk
[4]IMCS, University of Sindh, Jamshoro, Pakistan, fiazmemon@usindh.edu.pk
[5]IMCS, University of Sindh, Jamshoro, Pakistan, sm.chandio@usindh.edu.pk

ABSTRACT

In this proposed work we have been presenting an efficient Hardware and Software Controller for controlling appliances of home using combination of network technology and speech recognition technology. For speech recognition Linear Predictive Coding (LPC) has been implemented for extraction of features from acoustic inputs while states are designed using Hidden Markov Models at backend processing. The proposed model is designed to control local as well as remote home appliances using parallel ports of computer hardware interface.

Key words: HMMs, LPC, SR System

1. INTRODUCTION

Speech recognition refers to a computer's or software's ability to detect words and phrases in spoken language and transform them to a machine-readable format. The vocabulary of words and phrases recognized by rudimentary voice recognition software is limited, and it will only identify them if they are spoken clearly. Natural speech can be recognized by more complex software. Call routing, speech-to-text, voice dialing, and voice search are all speech recognition applications.

Controlling gadgets has always been a key topic of interest; much research has been conducted in this sector, and as a result, numerous methods of controlling the devices have been presented. To operate the equipment, most control devices have been developed utilizing both wired and wireless technologies, including remote control, internet, GSM, Wifi, and Bluetooth [1].

The problem with these proposed approaches is that they all require telecom technology in some form or another. Another issue we encountered was that the gadgets required physical help to function [2]. To address this quandary, we suggested an efficient and cost-effective solution that helps to address security issues by controlling devices using speech recognition.

The suggested system is extremely cost efficient since its hardware is not prohibitively expensive, making it accessible to all. Speaker recognition is the process of identifying and verifying voice waves based on information provided by the speaker.
We have built a hardware and software system which control Home Appliances Through Voice Recognition.

Overall, the goal of this project is to create a system that uses voice recognition technology to operate electronic appliances. Speech recognition technology is used in computer systems to transform spoken words into electronic text. This technology enables computer users to type text using their voices rather than keyboards. It is also a pretty natural and straightforward method of manipulating the simulation while keeping the user's hands free.

2. LITERATURE REVIEW

The two most popular techniques to speech recognition may be classified as "template matching" and "feature analysis." When applied correctly, template matching is the simplest approach with the best accuracy, but it also has the most restrictions. The initial stage, as with any technique to speech recognition, is for the user to say a word or phrase into a microphone. The electrical signal from the microphone is digitized and stored in memory using a "analog-to-digital (A/D) converter." The computer attempts to match the input with a digitized voice sample, or template, that has a known meaning to identify the "meaning" of the speech input. his method is similar to standard command inputs via a keyboard. The software provides the input template and uses a simple conditional statement to try to match it with the real input.

Because each person's voice is unique, the software cannot theoretically include a template for each conceivable user; hence, the program must first be "trained" using a new user's voice input before the computer can recognize that user's voice. During a training session, the program displays a printed word or phrase, and the user repeats it into a microphone several times. The software computes a statistical average of many samples of the same word and stores the

averaged sample as a template in a program data structure. The software's "vocabulary" is limited to the words or phrases spoken during the training session, and its user base is also restricted to individuals who have trained the program. This is known as a "speaker dependent" system. It can recognize up to 98 percent of words and phrases in vocabularies as little as a few hundred.

Feature analysis allows for a broader kind of speech recognition, which generally results in "speaker-independent" voice recognition. Rather than attempting to discover a perfect or near-precise match between the actual voice input and a previously saved speech template, this technique first analyses the voice input with "Fourier transforms" or "linear predictive coding (LPC)", The algorithm then looks for similarities between the predicted inputs and the actual digitized voice input. Because these commonalities will exist for a wide variety of speakers, the system does not need to be educated for each new user. Accents, variable pace of delivery, pitch, loudness, and intonation are examples of speech variations that the speaker-independent approach can manage but pattern matching cannot. The range of accents and inflections employed by speakers of different countries has proven to be one of the most difficult challenges in speaker-independent speech detection. The recognition accuracy of speaker-independent systems is somewhat lower than that of speaker-dependent systems, often ranging between 90 and 95 percent.

One voice recognition-based home appliance that could only handle one home appliance with sound was controlled [3,4,6 7]. The HM2007 chip was used to drive a wheelchair for disabled persons [5]. A large number of automobile manufacturers have used Arduino boards to provide diverse functionality [8, 9, 10, 11, 12, 13].

To the best of our knowledge, no such work has been done in the past using parallel port, Network Technology and Speech recognition technology using LPC and HMMs for controlling home appliances We presented an efficient and cost-effective solution for controlling household appliances.

3. PROPOSED MODEL

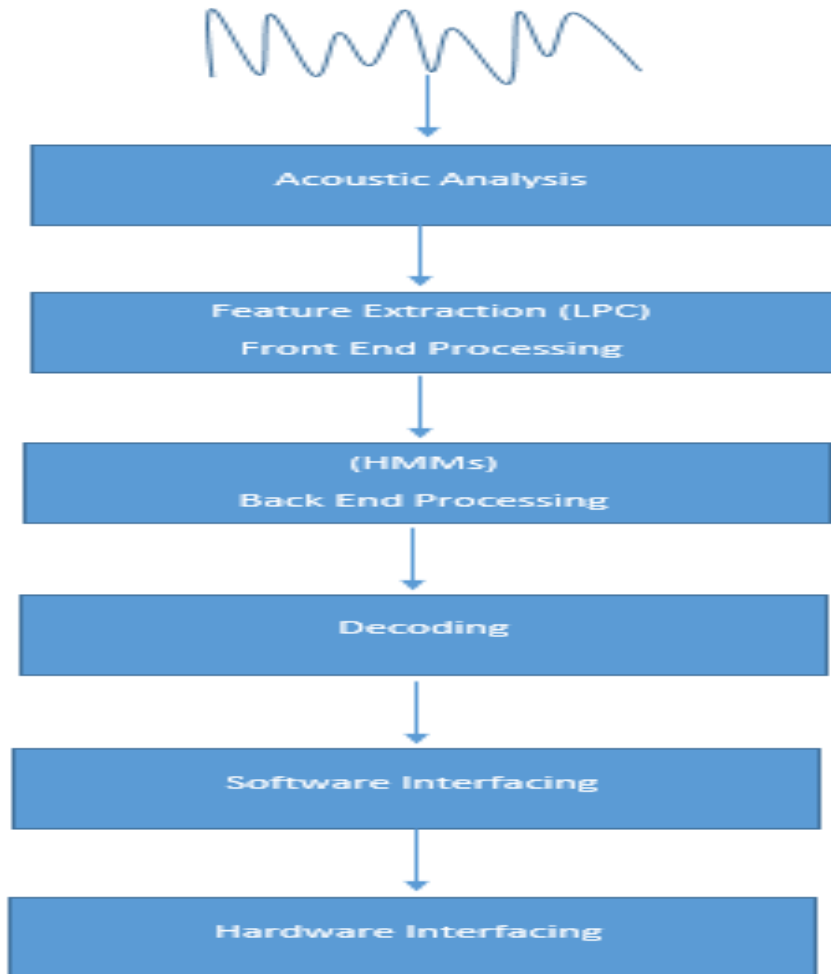Many number of steps are involved in this proposed model as shown in Figure 1,

Figure 1 : Proposed Model

first step is the user speech words or phonemes using MIC, list of words / Phrases included in this project are Table# 1:

Words / Phrases (On)	Words / Phrases (Off)
Bulb On	Bulb Off
Fan On	Fan Off
Tube Light On	Tube Light Off
Computer On	Computer Off
TV On	TV Off
Mobile On	Mobile Off
Motor On	Motor Off
Printer On	Printer Off
Open All	Off All

Table# 1 List of Phrases

In second step acoustic analysis is performed over spoken speech waveform, in third step LPC (Linear Predictive Coding) technique is applied to find out features from analyzed acoustic signals, in fourth step HMMs (Hidden Markov Models) is applied using word level modeling method which does internal transcription of speech. An example of (Bulb On and Bulb Off) is given in Figure 2.
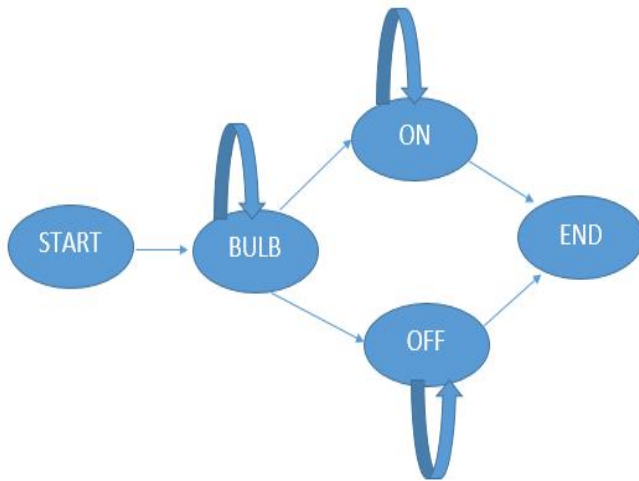

Figure 2: HMM States of (Bulb On & Bulb Off)

In fifth step decoding is performed to get transcription of speech, in the sixth step after searching and matching the required task is taken place to between hardware interface and software interface. In this system speech recognition processing takes place at user end while controlling of hardware takes place at system which was intended to control hardware (home appliances) connected to that system.

4. SYSTEM DEVELOPMENT

This project development consists of four parts. The first step is to create software. The graphical user interface (GUI) serves as the user's interface to the system. Visual Studio 2010 is used to create the GUI. This program includes the ability to hear voices and use the microphone.

Figure# 3 shows a software interface in which user could enter port# and IP to control hardware interface (Home appliances) of that system before connecting to appropriate system status of parallel port is (Off) while after connecting to appropriate system status of parallel port is (On) as shown in Figure 3.

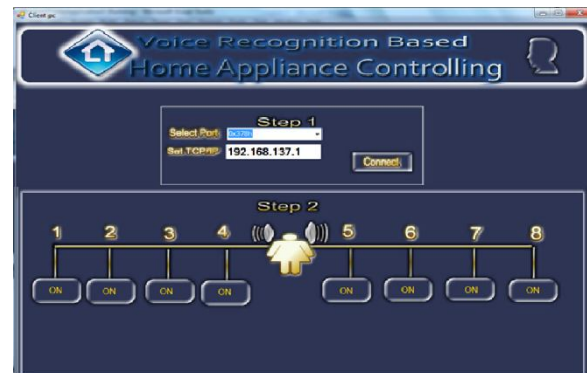Figure 3: (Selection of Port# and IPV4) (Off Status)

Figure 4: (Selection of Port# and IPV4) (On Status)

The Second part is used simple Socket programming to apply the benefit electronic text by bass to another pc through TCP/IP protocol then the PC performs third action. As illustrated in Figure 5, a server generally runs on a specific computer and has a server-socket that is bound to a certain port number. As seen in Figure 6, the server simply waits for a client to submit a connection request by listening to the socket. The client is aware of the hostname of the computer running the server as well as the port number to which the server is connected. The client attempts to connect to the server using the server's machine and port to begin a connection request. As seen in Figure 7, if everything works as intended, the server accepts the connection. Following acceptance, the server receives a new socket linked to a different port. To continue listening to the old socket for connection requests while catering to the needs of the connected client, a new socket (and hence a different port number) is required. If the connection is accepted, a socket is successfully created on the client side, and the client can now

communicate with the server via the socket. It should be noted that the client socket is unrelated to the port number used to connect to the server. Instead, the client is assigned a port number that is specific to the computer on which it is running. Both the client and the server may now communicate by writing to and reading from their respective sockets.
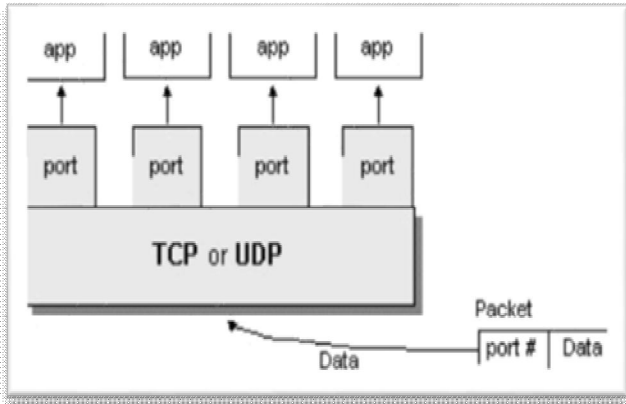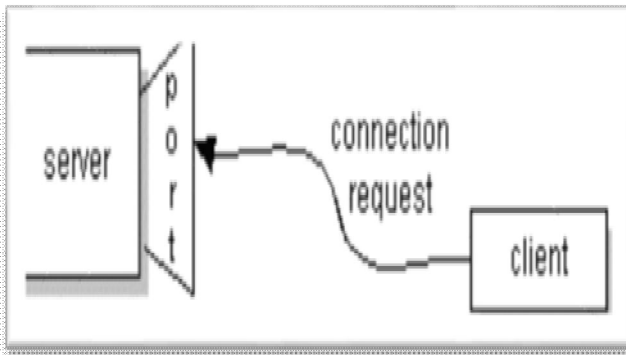

Figure 5: (Socket + TCP/IP)
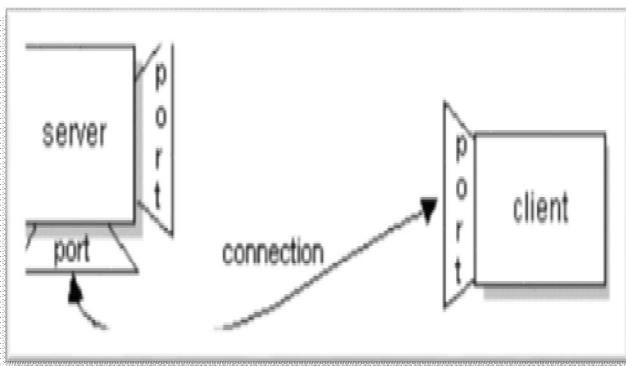

Figure 6: (Client Request for Port)

Figure 7: (Server Response of connection to client)

The Third part is using simple circuit as prototype to trigger the output. Figure# 8 shows the block diagram of controlling devices through speech recognition which elaborates the representation of simple components which are used such as bulb represents a lamp, motor represents a fan and relay represents a TV on. The Figure 9 shows the overview of system that defines how a user interface with system.
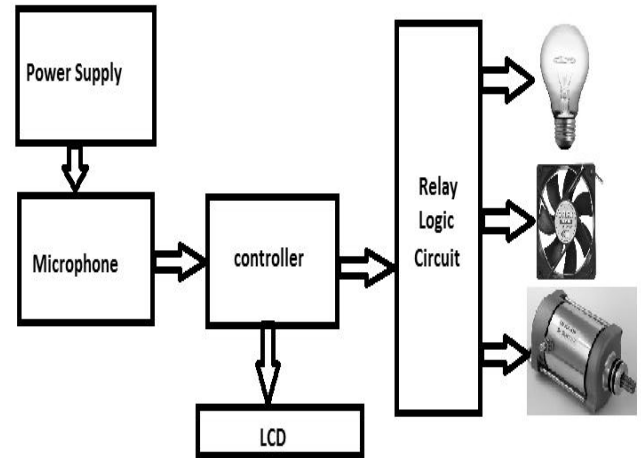

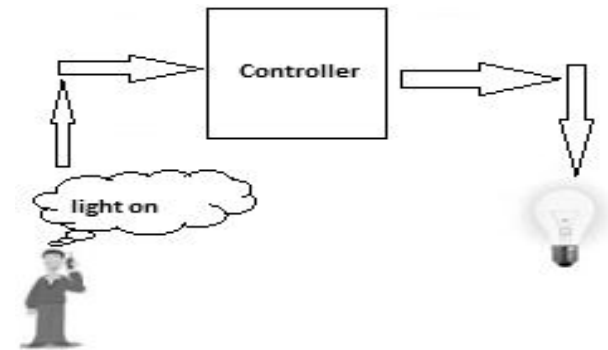Figure 8: Block Diagram of Controlling Devices through Speech Recognition

Figure 9: Overview of System

The fourth section involves using a parallel port to connect the circuits to the system (personal computer). This speech recognition technology system works by the user providing a vocal input, which the system subsequently accepts. The signal travels through the parallel port and activates the output shown in the circuit. Otherwise, the output will not activate and another input will be required. The interface circuit, in conjunction with the included software, may be used with any PC's printer connection to control up to eight devices. The interface circuit is designed for a single device, which is controlled by the D0 bit at pin 2 of the 25-pin parallel port. As illustrated in Figure# 10, identical circuits for the remaining data bits D1 through D7 (accessible at pins 3 through 9) must be similarly connected. The inclusion of an opto-coupler guarantees that the PC is completely isolated from the relay driver circuitry. Printers are still the most frequent devices attached to the port, but external tape and disc drives, as well as scanners, are also popular alternatives. A parallel-port-based network interface or joystick may be used by laptop computers. There are hundreds of parallel-port devices for use in data collecting, testing, and control systems for particular purposes. And the parallel port is the preferred interface for many one-of-a-kind and small-scale applications that require connection between a computer and an external

device. Despite its ubiquity, the parallel port has always been difficult to deal with. Several variants on the original port's design have evolved over the years, but there is no one source of documentation that explains the port in all of its incarnations.
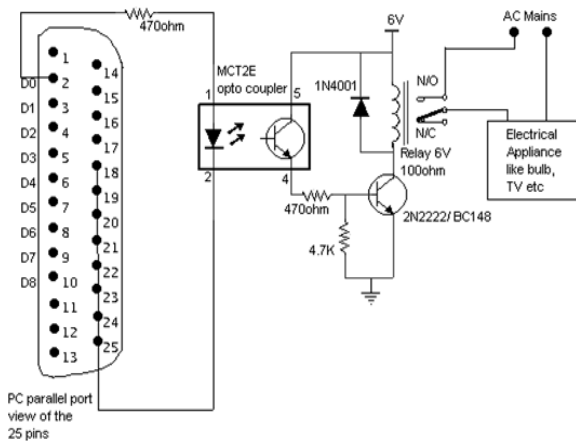

Figure 10: Parallel Port Interface

A computer based system is created to use as an interface to the speech recognition engine which integrated in Microsoft Windows. A software application is developed with graphical user interface (GUI) which designed using Microsoft Visual Studio 2010. Thus, it is easy to be used by everyone. From the software application, it sends an output signal to the hardware circuit through the parallel port. The output signal control the electronic appliances on or off.

5. SOURCE CODE

```
Initialize Speech Recognition
recognitionEngine = new
SpeechRecognitionEngine();
recognitionEngine.SetInputToDefaultAudioDevice();


Create Speech Recognition Grammar
    private Grammar
CreateSampleGrammar()
     {

  Choices TextChoices = new Choices(new
string[] {"Bulbeon","Bulbeoff", "Fanon",
"Fanoff",  "Tubeliteon", "Tubeliteoff",
"Computeron","Computeroff",
"TVon","TVoff", "Mobileon","Mobileoff",
"Motoron","Motoroff",
"Printeron","Printeroff",
"Openall","Offall"});
```

```
      GrammarBuilder grammarBuilder =
new GrammarBuilder();

grammarBuilder.Append(TextChoices);
      Grammar g = new
Grammar(grammarBuilder);
      g.Name = "Available programs";
      return g;
    }
```

Load grammar
```
GrammarBuilder grammarBuilder = new
GrammarBuilder();

grammarBuilder.Append(TextChoices);
            Grammar g = new
Grammar(grammarBuilder);
```

Register Speech Recognition
```
recognitionEngine.SpeechRecognized += (s, args) =>
```

Speech Recognition Event Handler
```
recognitionEngine = new
SpeechRecognitionEngine();
recognitionEngine.SetInputToDefaultAudioDevice();
recognitionEngine.SpeechRecognized += (s, args) =>
{
foreach (RecognizedWordUnit word in
args.Result.Words)
{

if (word.Confidence > 0.8f)
{

MessageWriter.Write(word.Text);
}
```

6. CONCLUSION

The main objective of this project was to design and implement the speech recognition technique to control the electronic and electrical appliances which can ease a person's daily activity. This makes the life easier and convenient. Nowadays, many applications can be reached without use of hands such as Bluetooth interact with mobile phone. In addition, it helps to reduce the utility expenses. Server that

recognized voice and Client receive text of voice then it gives signals to parallel port. hardware that are connected with parallel port produce volt signal to hardware. the hardware work on these instructions that denotes by server ("Fan on"" Bulb on" "Tube light on"" Computer on" "TV on" "Mobile on" "Motor on"" Printer on") every word open one parallel data port and ("Fan off" "Bulb off" "Tube light off" "Computer off" "TV off" "Mobile off" "Motor off" "Printer off") every word close one parallel data port. Also change label color and text on off and color green, Transparent.

This project is a network based computer application with speech recognition technology to control the electronic and electrical appliances such as fan, lamp, door lock, and the others. This application enables users to control any home appliances by inserting voice commands without the use of hands. This is useful for both able and disabled users. In addition, it is essential for people with limited physical ability.

REFERENCES:
[1] Nadeem Ahmed Kanasro, Mujeeb-U-Rehman Jamali, Ghulam Muhammad, Farhan Ali Surahio "Controlling of Home Appliances using Arduino UNO Board with Infrared (IR) and Radio Frequency (RF) Technologies" International Journal of Advanced Trends in Computer Science and Engineering, Volume 9, No. 5, September – October 2020 Online at
http://www.warse.org/IJATCSE/static/pdf/file/ijatcse317952020.pdf
https://doi.org/10.30534/ijatcse/2020/317952020

[2] M Z. RAZA, S. SOOMRO, M. S. EHSAN, A. G. MEMON, M. SIDDIQUE "Implementation of Device Control Technique Using Voice Signal" *Sindh Univ. Res. Jour. (Sci. Ser.) Vol.45 (4) 721-726 (2013)*

[3] Zhang H. and Wei Lieh Ng, (2010) Speech Recognition Interface Design for In-Vehicle System:
Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI 2010), November 11-12, 2010, Pittsburgh, Pennsylvania, USA

[4] Brouwer, W. H., W. Waterink,, P.C. Van Wolffelaar, and T. Rothengatter, Divided attention in experienced young and older drivers: Lane tracking and visual analysis in a dynamic driving simulator. Human Factors, 33 (5). 573 - 582.

[5] Choy M, "Voice Recognition Home Automation System". Degree. thesis. Electrical Engineering Department, University Teknologi Malaysia. 2010.

[6] Han Siong, J. (2009) Automated Home Lighting System, Degree. thesis. Faculty of Electrical Engineering, University. Technology Malaysia.

[7] Hashimah, I. (2004.) Motion Control using voice for wheel chairapplication. M.sc. thesis. Faculty of Electrical Eng., University Technology Malaysia.

[8] Joanne L and S., Lalande, (2005) Perofrming E-Mail Tasks While Driving: The Impact of Speech-Based Tasks on Visual Detection. in 3rd International Driving Symposium on Human Factors in Driver Assessment, (Rockport, Maine, 2005)

[9] Nakano, T., (2008) Flexible Shortcuts: Designing a New Speech User Interface for Command Execution. in CHI (Florence, Italy, 2008).

[10] Sonam K., and Kavita A., Controlling (2012) Of Device Through Voice Recognition Using Matlab: International Journal of Advanced Technology & Engineering Research (IJATER), Volume 2, Issue 2, March 2012.

[11] NADEEM AHMED. KANASRO, H.U. ABBASI, M.R. MAREE, A.G. MEMON, "Speech Recognition based web scripting from predefined Context Free Grammar (Language Model & Grammar) programmed in Visual Programming and Text Editor," Sindh Univ. Res. Jour. (Sci. Ser.) Vol.45 (3) 634-639, September- 2013

[12] NADEEM AHMED KANASRO, HABIBULLAH U.ABBASI, ABDUL GHAFOOR MEMON, MUJEEB-U-REHMAN MAREE, KAMRAN TAJ PATHAN. "Speech Recognition in context of predefined words, phrases and sentences stored in a database and its analysis, designing, development and implementation in an application". *International Journal of Advances in Computer Science and Technology Volume 2, No.12, December 2013, ISSN 2320 – 2602*. Impact Factor: 0.499

[13] NADEEM AHMED. KANASRO, NAJMA IMTIAZ ALI, GHULAM MUHAMMAD, MUJEEB U REHMAN MAREE, A.G MEMON "Finding Accuracy of Utterance of Language Model" International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 12 (2020), pp. 5129-5134 © International Research Publication House. http://www.irphouse.com
http://www.irphouse.com/volume/ijertv13n12.htm,
http://www.irphouse.com/ijert20/ijertv13n12_146.pdf