ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 12, March 2017

Enhancement and Effective Communication for Speech Impairment People

K.N.Baluprithviraj[1], N.Mahesh[2], K.NandhaKumar[3]
Assistant Professor, Department of EIE, Kongu Engineering College, Perundurai-638060, Erode, Tamil Nadu, India[1,2]
U.G. Student, Department of EIE, Kongu Engineering College, Perundurai-638060, Erode, Tamil Nadu, India[3]

Abstract – In this paper, a new method of Augmentative and Alternative Communication (AAC) device for people with severe speech impairment the Voice-Input Voice-Output Communication Aid (VIVOCA) is presented. The VIVOCA recognizes the disordered speech of the user and builds messages, which are converted into synthetic speech. A novel methodology for building small vocabulary, speaker-dependent automatic speech recognizers with condensed amounts of educating data, was applied in this paper. The output showed that this method is successful in generating good recognition performance on highly disordered speech, even when recognition perplexity is increased. This method of selecting message and building technique traded off various factors including construction speed of message and message outputs range available. The VIVOCA was evaluated in a field trial by individuals with moderate to severe Dysarthria and confirmed that they can make use of the device to produce intelligible speech output from disordered speech input.

Keywords- Voice-Input Voice-Output Communication aids, voice recognizer, voice synthesizer, microphone, speech recognizer**.**

I INTRODUCTION

Spoken language communication is a fundamental factor in quality of life, but as many as 1.3% of the population cannot use natural speech reliably to communicate, especially with strangers. For instance, the speech of people with moderate to severe dysarthria, the most common speech disorder affecting 170 per 100 000 of population, is usually unintelligible to unfamiliar communication partners. For these people, their speech impairment can preclude them from interacting in a manner that allows them to exploit their potential in education, employment and recreation. Speech impairment is frequently related with more physical disabilities as a result of progressive neurological conditions such as motor neurone disease, congenital conditions such as cerebral palsy, or acquired neurological conditions as a result of stroke or traumatic brain injury.

Recent technological tools for communication, Voice-Output Communication Aids, generally rely on a switch or keyboard for input. They do not readily make possible to normal communication as they are relatively slow and disrupt eye contact. The existing automatic speech recognition systems can work better for people with mild and even moderate dysarthria and it shows that there is an inverse relationship between the degree of impairment and the speech recognition accuracy.

II AUTOMATIC SPEECH RECOGNITION METHOD

Automatic Speech Recognition has been used by people for many years as a technique of access to technology with disabilities but unimpaired speech, it received little attention as a potential input channel for Voice Output Communication Aid.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 12, March 2017

Previous prototypes of Voice-Input Voice-Output Communication aids (VIVOCAs) have been reported, but have not been tested extensively with users or reached the stage of becoming available as commercial products. A wide range of user requirements were elicited and the VIVOCA was implemented to meet these requirements where feasible.

2.1 Description of Proposed Method

The block diagram of the Enhancement and Effective Communication of Speech Impairment People system is shown in Figure 1

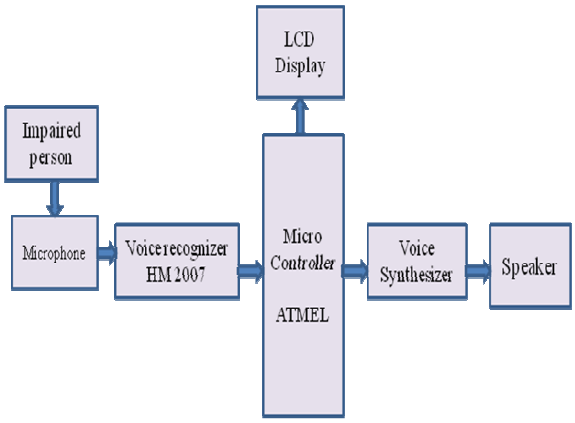

Figure 1: Block Diagram of the Enhancement and Effective Communication System

2.2 Microphone

. It is a device which converts the sound signal into an electrical signal. In this case Personal Digital Assistant (PDA) which has internal microphone is preferred for better input. The PDA acquire voice input from the user through a microphone, which can either be Bluetooth or wired or lapel-type, or the internal microphone of the PDA.

2.3 Voice Recogniser

The speech recognition system is a completely assembled and easy to use programmable speech recognition circuit. Programmable, in the sense that user trains the words (or vocal utterances) user wants the circuit to recognize. This board allows user to experiment with many facets of speech recognition technology. It has 8 bit data out which can be interfaced with any microcontroller for further development. The various interfacing applications which can be included for controlling home appliances, robotics movements, Speech Assisted technologies, Speech to text translation, etc. Speech recognition is divided into two broad processing categories; speaker dependent and speaker independent. The system which is speaker dependent is trained by the individual who will be using the system. Speaker independent is a system trained to respond to a word regardless of who speaks. Hence, the system has to respond for huge variety of speech patterns, inflections and enunciation's of the target word.

2.4 Microcontroller

The AT89C51 is a low-power, high-performance CMOS 8-bit microcomputer with 4 Kbytes of Flash Programmable and Erasable Read Only Memory (FPEROM). The Atmel AT89C51 is a powerful microcomputer which provides a highly flexible and cost effective solution to many embedded control applications by including an 8-bit versatile CPU with Flash on a monolithic chip. The AT89C51 provides the following standard features: 4Kbytes of Flash, 128 bytes of RAM, 32 I/O lines, two 16-bit timer/counters, five vector two-level interrupt architecture, a full duplex serial port, on-chip oscillator and clock circuitry. This device has ability to interface with TTL based devices. The 89C51 contain 34 general-purpose, or working, registers. Two of these, registers A and B hold results of many instructions, particularly for arithmetical and logical operations. The other 32 are arranged as part of internal RAM in four banks, Bank0-Bank3, of eight registers each. Program instruction bytes are fetched from locations in memory that are addressed by the PC. Program ROM may be on the chip at addresses 000h to FFFh, external to the chip for address that exceed FFFh, or totally external for all address from 0000h to FFFFh.

2.5 Full Duplex Serial Data Receiver/Transmitter

Computers must be able to communicate with other computers in modern multiprocessor distributed systems. One cost-effective way to communication is to send and receive data bits serially. The 89C51 has a serial data communication circuit that uses register SBUF to hold data. Register SCON controls data communication, register PCON controls data rates, and pins RXD (P3.0) and TXD

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 12, March 2017

(P3.1) connect to the serial data network. There are four programmable modes for serial data communication that are chosen by setting the SMX bits in SCON. Baud rates are determined by the mode chose.

2.6 Liquid Crystal Display

16x2 LCD represents that it can display 16 characters per line and there are 2 such lines. A command is an instruction given to LCD to do a predefined task like initializing it, clearing its screen, setting the cursor position, controlling display etc. The data register will have the data which is to be displayed on the LCD. The data is the ASCII value of the character to be displayed on the LCD.

2.7 Voice Synthesizer

Impaired person's voice is compared with the trainer's voice in the microcontroller and produces a data as an output signal and it is given to voice synthesizer. One of the user requirements for the system output was that it should be possible to have both pre-recorded and synthetic output, and that the synthetic output should be as natural sounding as possible

III METHODOLOGY OF THE PROPOSED SYSTEM

In this paper, microphone intakes disordered speech as input signal. The speech is processed by microphone and recognized by a speech recognizer. The recognized words are passed to a message building module. The schematic diagram of the speech impairment technique has shown in Figure 2.

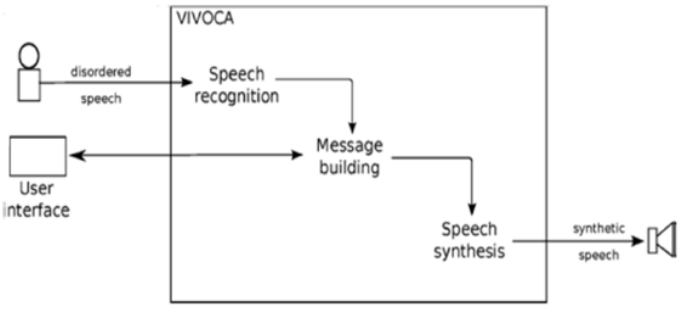

Figure 2: schematic diagram of the speech impairment technique

The message building module will update the screen based on this input, potentially supply audio feedback to the user, and determine the range of possible future inputs. This process continues in an iterative fashion as the user builds their message. When the message is complete it is passed to the speech synthesizer, producing intelligible spoken output via a speaker.

3.1 Voice Recognizer HM 2007 Circuit

HM 2007 is a single chip CMOS voice recogniser LSI circiut with the in-chip analog front end, voice analysis, recognisation process and system control functions. A 40 isolated –word voice recognition system can be composed of external microphone, keyboard, 64k SRAM and some other components.

3.2 Speech Recognition System

Speech recognition will become the method of choice for controlling appliances, toys, tools and computers. At its most basic level, speech controlled appliances and tools allow the user to perform parallel tasks i.e. hands, ears and eyes while working with the tool or appliance. The heart of the circuit is the HM2007 speech recognition IC. This speech recognition IC can recognize up to 20 words, each word having a length of 1.92 seconds.

Features

- Self-contained stand alone speech recognition circuit
- User programmable
- Up to 20 word vocabulary of duration two second each
- Multi-lingual
- Easily interfaced to control external circuits & appliances

The keypad and digital display are used to communicate with and program the HM 2007 chip. The keypad contains 12 number of normally open momentary contact switches. When the circuit is turned on, "00" is on the digital display, the red LED (READY) is lit and the circuit waits for a command.

3.3 Training Words for Recognition

Press "1" (display will show "01" and the LED will turn off) on the keypad, then press the TRAIN key (the LED will turn on) to place circuit in training mode, for word one. Say the target word into the

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 12, March 2017

onboard microphone (near LED) clearly. The voice input's signals are accepted and it is notified in the circuit by blinking the LED off then on. The word is now identified as the "01" word. If the LED did not flash, start over by pressing "1" and then "TRAIN" key. User may continue training new words in the circuit. Press "2" then try to train the second word and so on. The circuit will accept and recognize up to 20 words. It is not required to train all word spaces. If only require 10 target words that are all need to train.

Repeat a trained word into the microphone. The word count should be displayed on the LCD display. For instance, if the word "directory" was trained as word number 20, saying the word "directory" into the microphone will cause the number 20 to be displayed.

Error Codes:
The chip is having following error codes.
55 = word to long
66 = word to short
77 = no match

Trained words can easily be changed by overwriting the original word. For instances suppose word six was the word "Capital" and user want to change it to the word "State". Simply retrain the word space by pressing "6" then the TRAIN key and saying the word "State" into the microphone. If one wishes to erase the word without replacing it with another word press the word number, then press the CLR key. Word six is now erased.

3.3 The Voice with Stress & Excitement

Stress and excitement alters ones voice. This affects the accuracy of the circuit's recognition. For instance assume if a man sitting at a workbench and user program the target words like fire, left, right, forward, etc., into the circuit. Then use the circuit to control a flight simulator game, Doom or Duke Nukem. In the heat of the action user voice will sound much different than when user were sitting down relaxed and programming the circuit.

The speech recognition system is speaker dependent, meaning that the voice that trained the system has the highest recognition accuracy. But user can simulate independent speech recognition. To make the recognition system simulate speaker independence one uses more than one word space for each target word. Now we use four word spaces per target word. Therefore we obtain four different enunciations of each target word. The word spaces 01, 02, 03 and 04 are allocated to the first target word. We continue to do this for the remaining word space. For instance, the second target word will use the word spaces 05, 06, 07 and 08. We continue in this manner until all the words are programmed.

3.4 Speech recognition Technique and its method

Speech recognition is divided into two broad processing categories; speaker dependent and speaker independent. Speaker dependent systems are trained by the individual who will be using the system. The drawback to this approach is that the system only responds accurately only to the individual who trained the system. This is the most common approach employed in software for personal computers. Speaker independent is a system trained to respond to a word regardless of who speaks.

In addition to the speaker dependent/independent classification, speech recognition also contends with the style of speech it can recognize. They are three styles of speech: isolated, connected and continuous. Isolated: Words are spoken separately or isolated. This is the most common speech recognition system available today. The user must pause between each word or command spoken. Connected: This is a half way point between isolated word and continuous speech recognition. It permits users to speak multiple words. The HM2007 can be set up to identify words or phrases 1.92 seconds in length. This reduces the word recognition dictionary number to 20.

IV RESULTS AND DISCUSSION

Good recognition performance is achieved on highly disordered speech. The recognition perplexity has been increased. This paper relied in a user training phase in which user practiced speaking to the recognizer. It obtained a quality of speech for an impaired person for a normal communication. The Prototype Model for Speech Impairment People is shown in figure 3. The final prototype device was evaluated in a user trial designed to assess the process of configuring the device for a new user as well as

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 12, March 2017

the performance of the device in real communication situations.

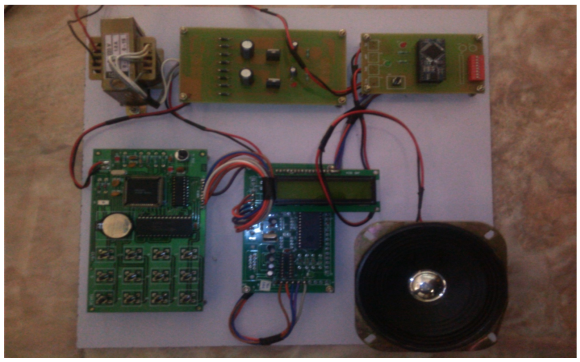

Figure 3: The Prototype Model for Speech Impairment People

CONCLUSION AND FUTURE SCOPE

In this paper, Voice Input Voice Output Communication Aid has been implemented mainly for speech impairment people. This paper made the impaired person to speak effectively. The previous method of implementation of the paper has lack of speed and efficiency. The VIVOCA has higher speed and accuracy. ZIGBEE module can be directly connected for home appliances. The other applications are voice speech controlled appliances and toys, computer assisted games, telephone assistance systems, voice recognition security and speech to speech translation. This design features that emerged from this process can make the final device more appropriate to a wider group of end users.

REFERENCES

1. Chaudhari,R., Deore.,R., and Gawali, B.(2012), 'P300 based Brain Computer Interface for Disabled', International Journal of Advanced Research in Computer Engineering and Technology(IJARCET) vol.1, Issue 6, pp.90-92

2. Falk.T.H, Chan.J, Duez.p, Teachman.G and Chau.T(2010), 'Augmenttive communication based on realtime vocal cord vibration detection', IEEE Trans. Neural System Rehabil. Eng., vol. 18, no.2, pp.159-163

3. Van der Meer.L, Sigafoos.J, O'Reilly.M.F, and Laniconi.G.E(2011),'Assessing preferences for AAC options in communication interventions for individuals with developmental disabilities: A review of the litereature', Res. Develop. Disabilites, vol.32, no.5, pp.1422-1431

4. Ben-Mosbah, B.(2006). Speech recognition for disabilities people. In Proceedings of the second information and communication technologies(ICTTA 2006) pp.864.869

5. B. O'Keefe, N. Kozak, and R. Schuller(2007) 'Research priorities in augmentative and alternative communication as identified by people who use AAC and their facilitators', Augmentative Alternative Commun.

6. D. Beukelman and P. Mirenda (2009), 'Augmentative and Alternative Communication', Paul H. Brookes.

7. L. J. Ferrier, H. C. Shane, H. F. Ballard, T. Carpenter, and A. Benoit (1995), 'Dysarthric speakers' intelligibility and speech characteristics in relation to computer speech recognition'

8. M. S. Hawley et al. (2010), 'A speech-controlled environmental control system for people with severe dysarthria'.

9. P. D. Green, J. Carmichael, A. Hatzis, P. Enderby, M. S. Hawley and M. Parker(2010), 'Automatic speech recognition with sparse training data for dysarthric speakers'

10. J U. Sandler and Y. Sonnenblick (2006), 'A system for recognition and translation of the speech of handicapped individuals'.