Journal of Information Engineering and Applications
ISSN 2224-5782 (print) ISSN 2225-0506 (online)
Vol.4, No.3, 2014
www.iiste.org

Computer Based Test using Speech Recognition System

AFOLABI A.O (Corresponding Author)
Department of Computer Science and Engineering
Ladoke Akintola University of Technology, Ogbomoso. Nigeria
E-mail: adefolabius@yahoo.com
ADIGUN A.
Department of Computer Science and Engineering
Ladoke Akintola University of Technology, Ogbomoso. Nigeria

Abstract
Speech recognition system is a system that can converts audio signals into machine-readable format to provide interaction with the system. In this paper the speech recognition system is applied within the E –Learning system. A headset microphone is interfaced to the computer system for speech production, a speech engine to help analyze each spoken text as programmed in the source code using XML format as the back end for stored text. This technique process data capturing more promptly than the existing process by keyboards or mouse, and can be used in working environment where these devices cannot be used and also for the disable learners that cannot use their hands or communicate through textual medium. Speech Recognition System is a developing technology and its application in this context is done within the Computer-Based Test as a pilot study. The speech SDK (Software Development Kits) used has an inbuilt speech recognition engine and inbuilt simulator, the engine provides the speech recognizer that is able to recognize the user's speech and integrate it with the inbuilt simulator.
Keywords: Speech Recognition System, Computer Based Test, E-Learning System

1. Introduction

Most speech recognition research, up to the 1980's, considered the major research problem to be one of converting a speech waveform (as an acoustic realization of a linguistic event) into words (as a best-decoded sequence of linguistic units). Many researchers also believed that the speech-to-text process was the necessary first step in the process that enabled a machine to be able to understand and properly respond to human speech. In field evaluations of speech recognition and understanding technology for a range of tasks, one important thing were learned about the speech communication process between humans and machines is that potential users of a speech recognition system tended to speak natural sentences that often did not fully satisfy the grammatical constraints of the recognizer (e.g., by including out-of-vocabulary (OOV) words, non-grammatical constructs, ill-formed sentences, etc.), and the spoken utterances were also often corrupted by linguistically irrelevant "noise" components such as ambient noise, extraneous acoustic sounds, interfering speech, etc.

Speech is a complex phenomenon. People rarely understand how it is produced and perceived. In human brain, thoughts are constructed into sentences and the nerves control the shape of the vocal tract (jaws, tongue, mouth, vocal cords etc.) to produce the desired sound. The sound comes out in phonemes which are the building blocks of speech. Each phoneme resonates at a fundamental frequency and harmonics of it and thus has high energy at those frequencies. The first three harmonics have significantly high energy levels and are known as formant frequencies. Each phoneme has a unique fundamental frequency and hence unique formant frequencies and it is this feature that enables the identification of each phoneme at the recognition stage. In general, speech recognition systems have stored reference templates of phonemes or words with which input speech is compared and the closest word or phoneme is given out.

2. DESIGN AND METHODOLOGY

Components Used

The speech recognition system is divided into six major components, as illustrated below. The components are:

i. Microphone Headset
ii. Audio Capture Subsystem
iii. Speech Recognizer
iv. User Interface
v. Inbuilt Simulator
vi. Speech Recognition System Evaluator
vii. Speech SDK
viii. Visual studio 2010

The headset is mainly for voice capturing. It interfaces with the audio capture subsystem existing on the machine. The Speech recognizer is an instance of the speech recognition engine gotten from the speech SDK. It does

actual speech acceptance and rejection. The various kits of the speech recognition have been simulated to support both multiple and single users via the recognizer mode enumeration class. The actual development was done in the visual studio 2010 environment which support c# programming language and the extensible markup language.

3. RESULTS AND DISCUSSION

The study imports the use of an Extensible Markup Language where words to be recognized by speech are programmatically collated. The XML constitute the backend while the interface was developed using a Windows Presentation Foundation application, it constitute the frontend where the user via his own speech answers certain questions.The following are the documentation of the analysis and possible outcome

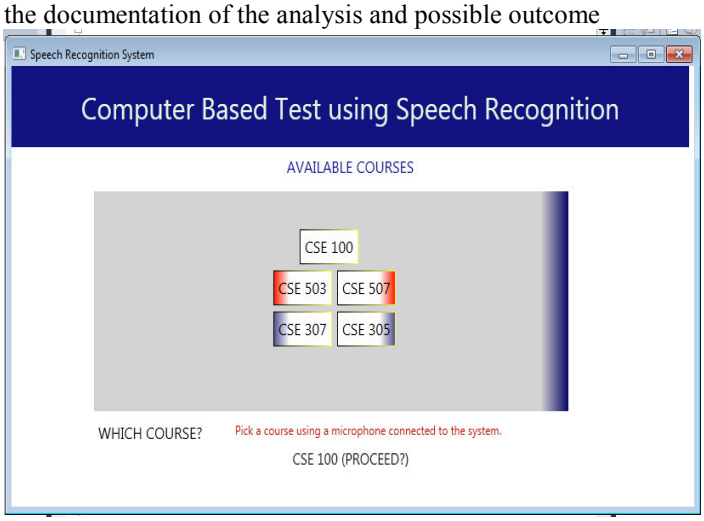

Figure 1 OPTION BY SPEECH

3.1 Option by Speech Implementation

In the above is an interface where the user picks up a course via speech. By say “CSE100”. After he says CSE100 comes up a proceed statement requesting the user to say “PROCEED” in order to access the second interface shown below.

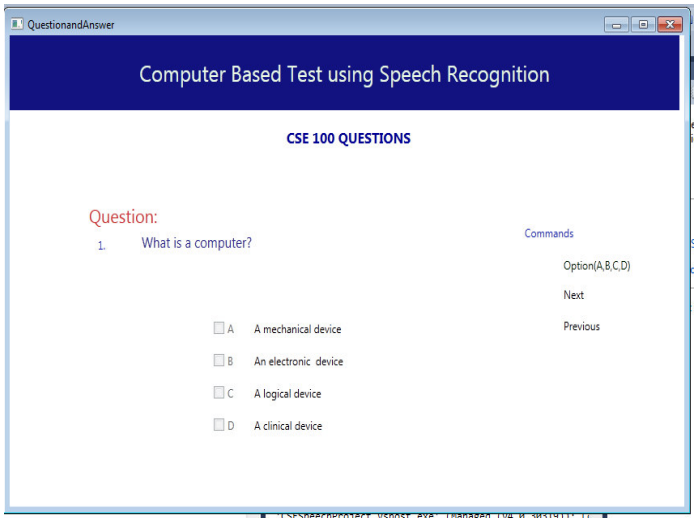

Figure 2 Second Interface

3.2 Second Interface Implementation

Immediately he says ‘PROCEED’ the second interface pops a up an interface with question and options let’s say options A-D. Any option he chooses would be checked via speech. Assume he picks option A as shown below.

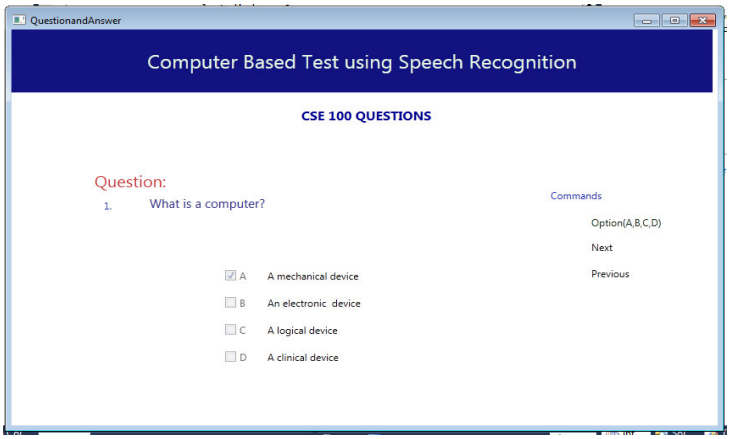

Figure 3 CBT Questions Interface

Now that the option A has been checked, he can then proceed by saying "NEXT". After that comes up the figure below

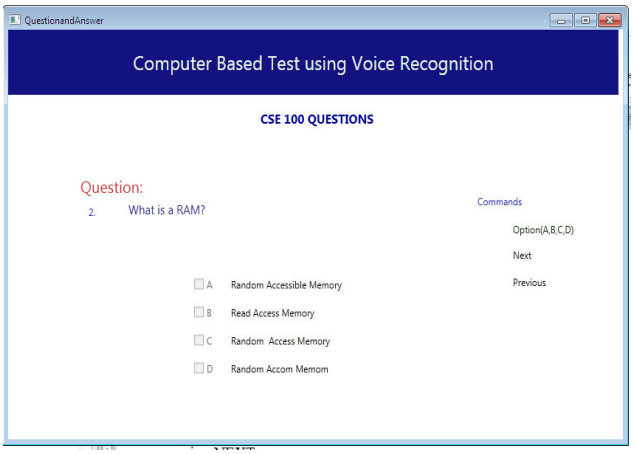

Figure 4 Option Interface

Implementation

In the third interface is a second question where the user can pick an option again and either "Next" or go "Previous". The user picks 'C' and says "Next" as shown below

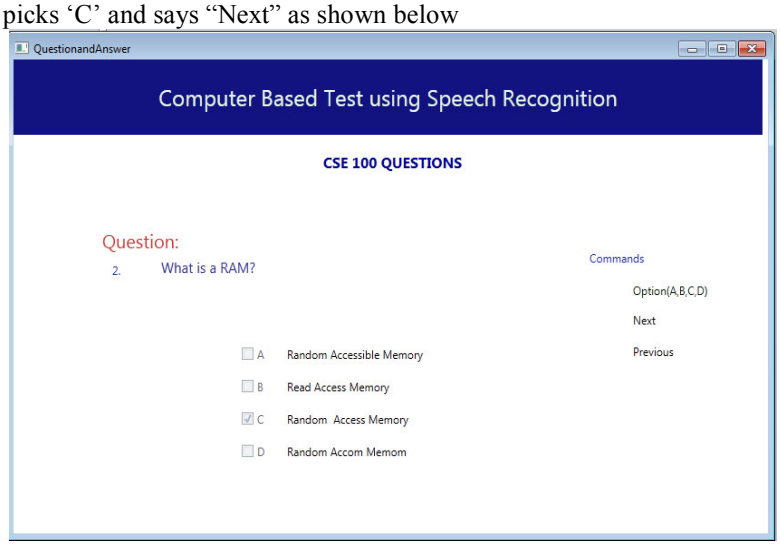

Figure 5 Third Interface

He says "Next" to pop up the Fourth interface as shown below.

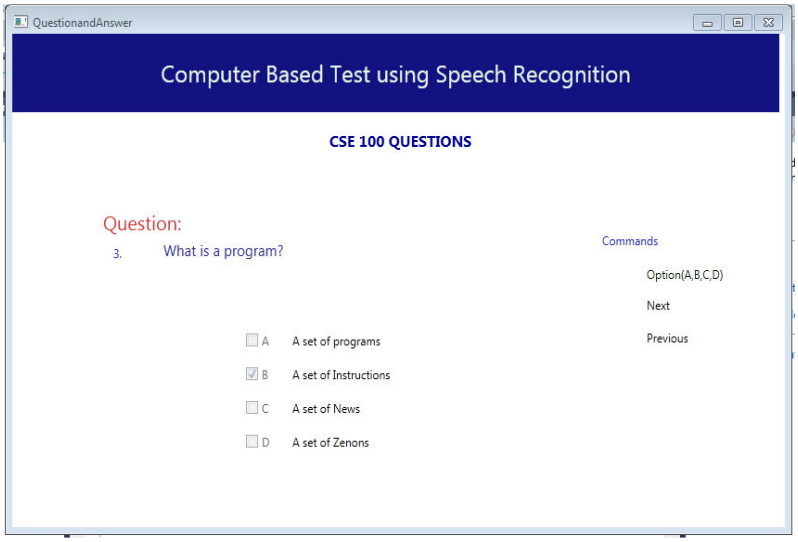

Figure 6 Questions Interface

Questions Interface Implementation

In the fourth interface the user picks "B" and says "Next" to pop up the fifth interface which carries 4th question with "C" as the checked answer as shown below.

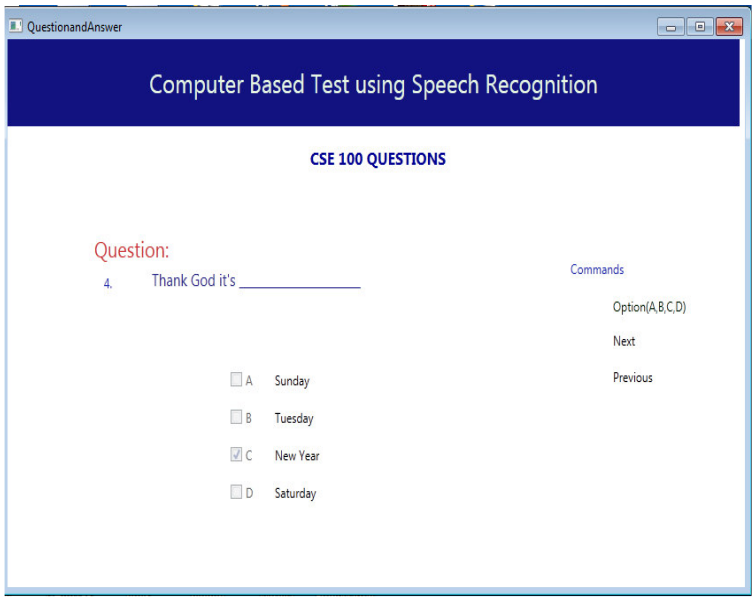

Figure 7 Fifth Interface

In the Fifth interface he says "Next" again to access the sixth interface. This carries the 5th question with option "C" as the checked answer as shown below.

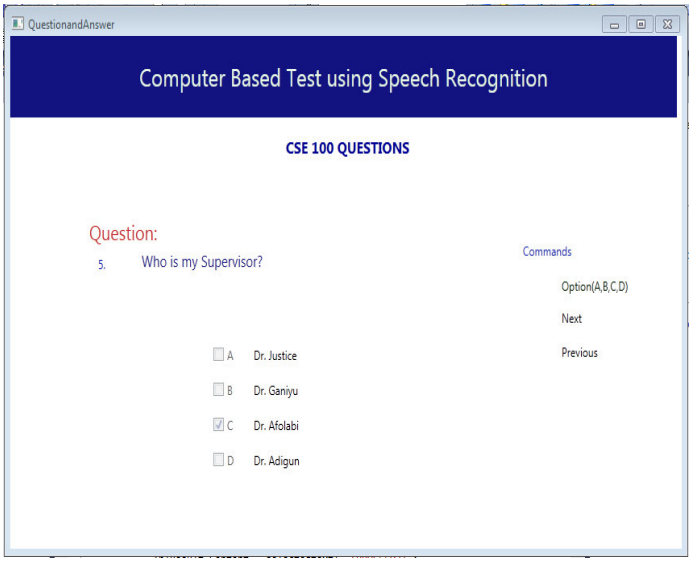

Figure 8 Sixth Interface

Sixth Interface Implementation

In the sixth interface is the sixth question with option "D" the checked answer as shown below.

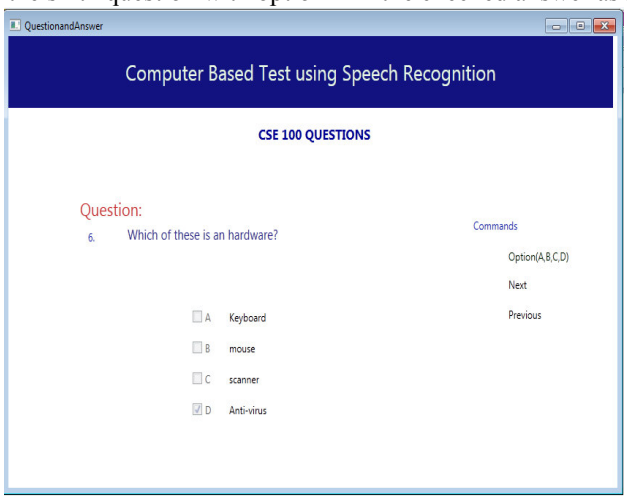

Figure 9 Seventh Interface

Seventh Interface Implementation

On answering the question he says "Next" so that the total score is collated and displayed as shown below.

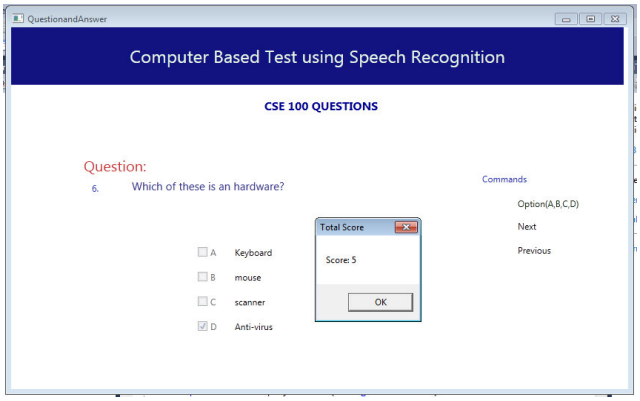

Figure 10 Final Interface

Final Interface Implementation

The final interface implementation shows the total score of the checked answers picked by the user's speech. Thus the Computer-Based Test Application was made possible via speech recognition.

Performance Analysis

It was however discovered that the performance of the application is affected by a noisy environment. Under a very noisy environment the application gives no response to the user's voice. Under a slight noisy environment the response is under a probability of 0.5 but in a very quiet environment the probability is just 1. Also it was discovered that the option 'A' responds under 'ha' sound and sometimes under 'hay' sound. The complication was as a result of natural variations in spoken speech that are difficult to handle in discrete systems. Examples include variations in intonation and accent (acoustic level), and variations in meaning such as the difference between "there" and "their" (linguistic level)

4 CONCLUSION

Speech recognition technique can process data capturing more promptly than the existing process by keyboards or mouse, and can be used in working environment where these devices cannot be used. In addition, since it is easy to move and control the device, and does not require repetitive key input and mouse works, it can remove the monotonous and tedious process of entering data. Speech Recognition System is a developing technology and it's applicable to e-learning. The Computer-Based Test Speech software is an application that support e-learning via user's speech. The speech SDK (Software Development Kits) used has an inbuilt speech recognition engine and inbuilt simulator. The engine provides the speech recognizer that is able to recognize the user's speech and integrate it with the inbuilt simulator. The software is useful for the disable who might also be willing to partake in the e-learning process. Some students and adults have physical disabilities that preclude their using a standard keyboard or mouse effectively. For these students, speech recognition is one of several alternative input methods to be explored. Speech recognition may provide a more efficient means of controlling a computer that is less physically and cognitively taxing than other alternative input methods. Speech recognition technology can benefit students who have learning disabilities that interfere with their ability to spell and write. While many such students benefit from standard word processing, the visual-motor demands of keyboarding can be a major stumbling block that compounds the writing difficulties. Similarly students who are the poorest spellers are frequently unable to effectively use standard spell checkers. For whatever reason, if students' oral language skills far outstrip their ability to generate text with pencil and paper or standard word processing, speech recognition may enable them to become accomplished writers by circumventing the most frustrating aspects of text generation.

REFERENCES

1. John K. and Alan W.B., (2003): "CMU ARCTIC database for speech synthesis ".CMU-LTI vol. 3 pp.177.
Language Technologies Institute, School of computer science, Cranegie Mellon University
2. Husni M.,Moustafa E., Mansour A.,(2001):"Technique for high quality Arabic Speech synthesis". An international journal on information Sciences. June 28, 2001. Views the demonstration at www.elseview.com/locate/ins
3. Ken Fujisawa and Nick Campbell (1998):"Prosody _Based unit selection for Japanese Speech synthesis" The third ESCA/COCOSDA workshop (ETRW) on speech synthesis Jenolaon cave House, Blue
Mountain NSW, Australia,
4. Nicolas, pierre L., Mathiew A., "Towards improved HMM-Based Speech Synthesis using high level syntactical features"http://alpage.iniria.fr/alpc.en.html.
5. Odejobi O.A., Beaumont A.J.,Wong. (2006): "intonation contour realization for standard
Yoruba text to speech synthesis A Fuzzy computational approach, computer speech
6. Sadaoki Furui (2002) "Speaker Recognition", July 12, 2005, http://cslu.cse.ogi.edu/HLTsurvey/ch1node9.html
8. Markowitz, J. (2002), "Speech Recognition & Speaker Biometrics", August 4, 2005 http://www.jmarkowitz.com/glossary.html#anchor_glosstop
9. The Speaker Recognition Homepage (2002) "Speaker Recognition – Algorithms" June 16, 2005 from http://www.speaker-recognition.org/navAlg.html