INTELLIGENT VOICE-BASED E-EDUCATION SYSTEM: A FRAMEWORK AND EVALUATION

Ambrose Azeta, Charles K. Ayo, Aderemi Aaron Anthony Atayero, Nicholas Ikhu-Omoregbe

Covenant University, Ota, Ogun state, Nigeria,
azetaambrose@gmail.com, ckayome@yahoo.com, atayero@ieee.org, nomoregbe@gmail.com

Abstract: *Voice-based web e-Education is a technology-supported learning paradigm that allows phone-access of learners to e-Learning web-based applications. These applications are designed mainly for the visually impaired. They are however lacking in attributes of adaptive and reusable learning objects, which are emerging requirements for applications in these domain. This paper presents a framework for developing intelligent voice-based applications in the context of e-Education. The framework presented supports intelligent components such as adaptation and recommendation services. A prototype Intelligent Voice-based E-Education System (iVEES) was developed and subjected to test by visually impaired users. A usability study was carried out using the International Standard Organization's (ISO) 9241-11 specification to determine the level of effectiveness, efficiency and user satisfaction. Report of our findings shows that the application is of immense benefit, based on the system's inherent capacity for taking autonomous decision that are capable of adapting to users' requests.*

Keywords: *Adaptation, Case-Base Reasoning, e-Education Framework, Intelligent, Ontology, VoiceXML*

1. INTRODUCTION

E-education entails the use of modern information and communication technology (ICT) in teaching and learning to enhance traditional classroom education and achieve better results in distance education, web-based learning and mobile learning [1]. Over the last decade, there has been a change of focus from web-based learning systems that merely turn pages of content to systems which present learning materials in such a way as to satisfy the needs of learners. This change of focus is especially important in modern learning methods, which place strong emphasis on learners' previous knowledge. The uniqueness of a learner is met by making the web-based learning content adaptive [2].The increasing number of Learning Management Systems (LMS) for online teaching, quiz, assignment delivery, discussion forum, email, chat, et cetera, means that dynamic educational online services will be needed for efficient management of all educational resources on the web. Selecting and organizing learning resources based on learner's interested is cumbersome [3]. The process of selection may be easier with the normal users, but for certain category of learners notably the visually impaired, navigating a Voice User Interface (VUI) for the desired learning content is a strenuous task.

Web-deployed VUI applications for educational purposes provide user accessibility to content via telephone. One of the tools used for developing these applications is Voice eXtensible Mark-up Language (VoiceXML). VUI applications are primarily developed to cater for the visually impaired [4]. The major problem is that the services of reusable learning objects currently available to regular users are not available for the visually impaired [5]. The voice-based web applications mainly designed for the visually impaired lack sufficient attribute of adaptable and reusable learning objects, which is a major requirement for this category of users as a result of their physical impairment. The goal of adaptive voice-based web learning is to adjust the content of the learning objects to suite user's knowledge level, whereas the recommendation services provide the most appropriate learning objects to users. The inability of existing e-Education voice applications to meet these requirements has some far reaching implications such as limited accessibility for certain users especially the visually impaired, and usability issues as a result of lack of features for reasoning, adaptation and recommendation. A significant contribution would be to introduce the concept of reusing learner's previous experiences which is often neglected in existing voice-based e-Education systems, to make them more adaptable and provide recommendation services to users' needs. Thus, we

have to take into account that the existing reference guidelines for developing these voice-based e-Education applications, although important, lack intelligent component services to approach the problem. A way of enhancing existing voice-based e-Education applications is by adapting their content to the needs of each student. It is a common belief that the performance of spoken dialogue system's could be enhanced through adaptation [6], and that adoption of artificial intelligence methodologies using case-based reasoning (CBR) to provide recommendation services to the different needs of students is essential[7], since CBR is the process of solving new problems based on the solutions of similar past problems.

From the foregone, it becomes obvious that there exists a need for providing a framework for improving the level of intelligence of voice-based e-Education applications. This paper aims at addressing the need as identified, and employing the framework so obtained for the development of a prototype Intelligent Voice-based E-Education System *(iVEES)* capable of improving learning processes using telephone and web-based technologies. The prototype application developed was tested in a school for the blind, and the result of evaluation reported. The proposed framework suffices as a reference model for implementing intelligent voice-based e-Education applications. Applications developed based on this framework would exhibit the necessary attributes of CBR, including adaptation and recommendation during interaction resulting in intelligence such as the ability of the system to take autonomous decision that will adapt to learners' request based on requirements. The application will therefore be helpful for people with physical access difficulties engendered by their impairment.

The research methodology adopted in this study involves a combination of CBR and Porter's stemming algorithm to realize the intelligent search agent for answers to tutorial questions. CBR is used because we desire a system that can reason from experience to save the time of teachers attending to large number of students. A system that will not only provide solution but recommend a viable solution is most appropriate for this research. The reason for using Porter Stemming Algorithm is to ensure that the suffix stripping process reduces the total number of terms in the information retrieval (IR) system, and hence reduce the size and complexity of the data in the system, which is always advantageous. A smaller data size results in a saving of storage space and processing time.

The remaining part of this paper is structured as follows: Section two discusses related work. Section three presents the proposed voice-based e-Education framework with specific emphasis on its adaptation

and recommendation component services. In section four, a description of an implementation and deployment architecture of the proposed system is presented. Section five reports the system evaluation and the paper is concluded in section six.

2. RELATED WORK

Learning contents are increasingly stored digitally and can be accessed in voice response form. This has opened up new opportunities for voice applications in the domain of e-Education. A voice-enabled web-based absentee system was developed in [8] on TellMe voice portal [9] to include record keeping of absentee calls from students, faculty and University staff. It was tested by a class of software engineering students. A student intending to miss a class called the VoiceXML telephone number and was led through an automated dialog to have his/her data, date and time of the call, courseID, and date of intended absenteeism recorded in a database. The system provided the instructors and other administrators' permanent record of absentees that can be accessed and displayed in various forms through a web interface. The researchers in [10] explored the integration of speech recognition technologies into m-Learning applications to reduce access barriers. An educational online forum accessible through mobile devices was built, based on an m-Learning framework proposed in [11]. In the same way, the result of their experience were presented in [12] with developing an experimental prototype for speech-based literacy e-Learning and application architecture suitable for literacy based e-Learning. The authors in [4] designed, implemented and deployed a voice-enabled application called V-HELP system where a portion of the Computer Science and Engineering (CSE) department website is voice-enabled using VoiceXML allowing the access of visually impaired student population.

These applications were developed using VoiceXML to handle record keeping of learners' data including data input and output through voice response. However, they are incapable of taking independent decisions apart from responding to what the user has requested from the system. The features of content adaptation and recommendation during information retrieval (IR) were not considered as an integral part of the applications. Thus, the framework used by these applications lack the required intelligence for reusable learning objects and this is the motivation for the proposed framework.

3. THE VOICE-BASED E-EDUCATION FRAMEWORK

A user-centric design approach is adopted in order to develop a framework that is based on voice

interaction. A survey was conducted to find out the requirements of an e-Education systems, and in particular that of the visually impaired. From the data collated from the survey questionnaire, significant differences were found between expert users, beginners and intermediate users. Expert users can do most of the typical tasks that users would normally do on the Voice User Interface (VUI) such as voice learning, participating in tutorials and checking examination results. As was expected, beginner users have little or no knowledge of some particular learning content. Based on the information gathered, the following points were identified as major requirements for the framework and application: i) Ability to provide users with services involving course registration, voice lecture, tutorial, examination and result; ii) ability to accommodate different types of users based on their respective learning profile; iii) ability to differentiate the course contents into different levels such as expert, intermediate and beginners; and iv) selection of different course contents based on users' profile. The proposed framework for iVEES is shown in Fig.1. The framework comprises of interaction and intelligent layer.
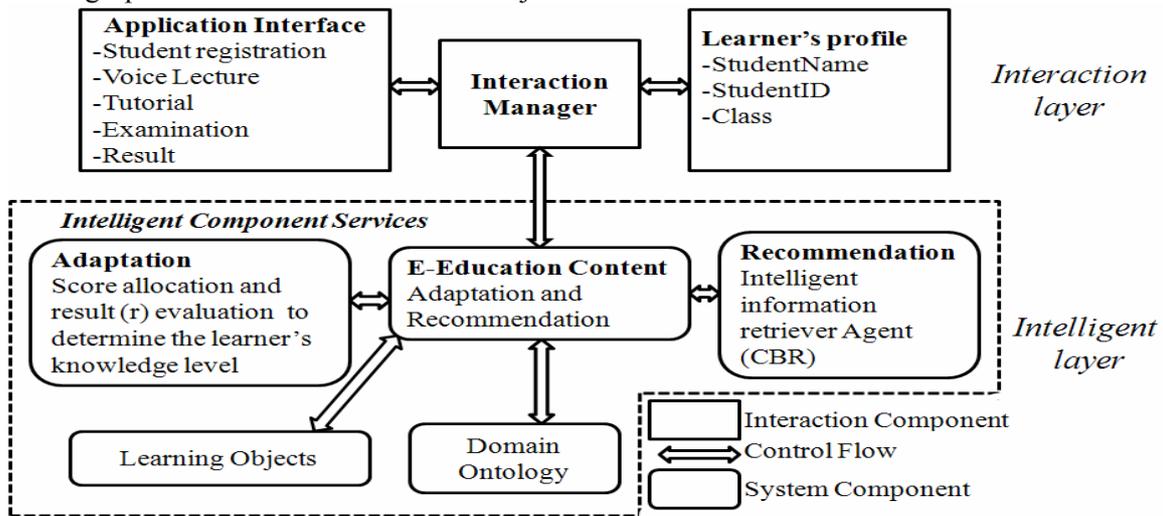

Fig. 1 – Intelligent voice-based e-Education framework

Learner's knowledge level is subdivided into three categories: beginner, intermediate and expert. A learner's profile determines whether he/she should receive beginner, intermediate or expert content of lecture, tutorial, examination and result module. The system navigation process is contained in the interaction manager. The classification as beginner, intermediate or expert level is done through content adaptation and recommendation using score allocation and result evaluation services. The intelligent information retrieval (IR) involving CBR and stemming is engaged at the recommendation services to provide recommended answers to tutorial questions using previous experience of learners, and also to expand the search for answers using stemming algorithm. Learning object consist of chunk of course material in text format which allow information to be presented in several ways to users. Tutorial questions asked by previous learners and e-Education data are stored in case knowledge and domain ontology respectively.

The adaptation component service uses the auto score allocation and result (r) evaluation model to determine the learner's knowledge level. The model is expected to be able to create the most suitable course content for each learner and control the passage from one knowledge level to another. There is within a particular knowledge level, an activity (quiz) containing at least one question. Before moving from one knowledge level to another, the system must evaluate the learner's performance through set of evaluations. The evaluation criteria are represented in Fig. 2.
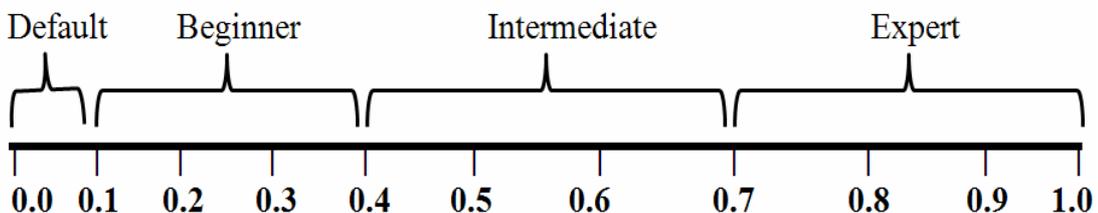

Fig. 2 – A model of the score evaluation criterion

Score allocated for each question:

$$Q = \sum_{k=1}^{q} Q_k \quad \text{and} \quad S = \frac{Max(r)}{Q} \qquad (1)$$

Evaluation result for each Activity:

$$Y = \left(\sum_{j=1}^{s} S_j \right) - W \qquad (2)$$

Evaluation result for each knowledge level:

$$R = \sum_{m=1}^{y} Y_m \qquad (3)$$

The following assumptions were made in formulating equations 1, 2 and 3.

r = Expected result for score evaluation criteria (see Fig. 2).

$Max(r)$ = The largest value of r.

R = evaluation result for each knowledge level t where t = 1, 2.., r.

Q = total number of questions for each activity k where k = 1, 2, ..., q.

S = score for each question j where j = 1, 2, ..., s.

Y = evaluation result for each activity m where m = 1, 2, ..., y.

W = weight for each knowledge level x where x = 1, 2, …, r. W is the initial value assigned to each knowledge level.

To be allowed to move from one knowledge level to another, the learner's result must satisfy the following transition criteria: $0.0 < =R < 0.1$(default), $0.1 < R < 0.4$ (beginner level), $0.4 < R < 0.7$ (intermediate level), and $0.7 < R <= 1.0$ (expert level).

The learner's experiences and situation were captured in the learner's profile. By using this experience, the system is able to offer to the current learner the best suited learning content. This experience was captured and provided using CBR. The recommendation services is designed using CBR, stemming and domain ontology. Our proposed system uses *Biology* domain ontology that represents specific knowledge, i.e., relationship between words used in Biology subject. The text tokenizer and porter stemming algorithm decomposes the whole textual information into sentences, and then into individual words with their stem for ease of retrieval of its synonyms. A Computation of the weight for every word and enhancement of term using ontology are required to improve retrieval effectiveness due to the huge amount of words.

Vector space models (VSM) is employed for performing query retrieval and assigning weights to the words in the query [13]. Two main components of the weight are term frequency and index document frequency (*TF/IDF*). *TF/IDF* estimates the importance of a term in a given document by multiplying the new term frequency (TF) of the term in a query by the term's inverse document frequency (IDF) weight. Weight of the ith term in j^{th} query is computed using the formula in equation 4[14]:

$$w_{i,j} = f_{i,j} * \log\left(\frac{n}{df_i} \right) \qquad (4)$$

The weighting terms $W_{i,j}$ method calculates the weight of each term or word from the stored cases and the inputted user's query to perform further matching. Where $f_{i,j}$ is the frequency of the ith term in j^{th} query. df_i is the number of queries that contain the word i. n is the total number of queries. To find the textual similarity between a stored case vector and a new case query vector, we apply cosine similarity function [15] for the textual information.

4. IMPLEMENTATION

A prototype application has been developed using Hypertext Preprocessor (PHP) for Web User Interface (WUI), VoiceXML for Voice User Interface (VUI), Apache for middleware and MySQL as database. The components of the application comprising design, implementation and development has been reported in [16, 17]. The WUI of the application presented in Fig. 3 contain screen shot of a sample tutorial question used to demonstrate recommendation component service of the application. The WUI is developed for system administration purposes such as upload of tutorial questions and lecture notes. Voxeo speech platform [18] was engaged as the speech server while CBR and porter's stemming algorithm were used to provide intelligent services. The layered architecture presented in Fig. 4 is the generic platform for the proposed framework. It shows the location of each component in the platform and consists of the presentation tier, logic tier and data tier. The database is separated from the client by the logic-tier. In the presentation tier, users are able to connect to the e-Education application. The components of the clients' interface are i) Personal Computers (PC) for the WUI, and ii) Land phone and Mobile phone for the VUI.
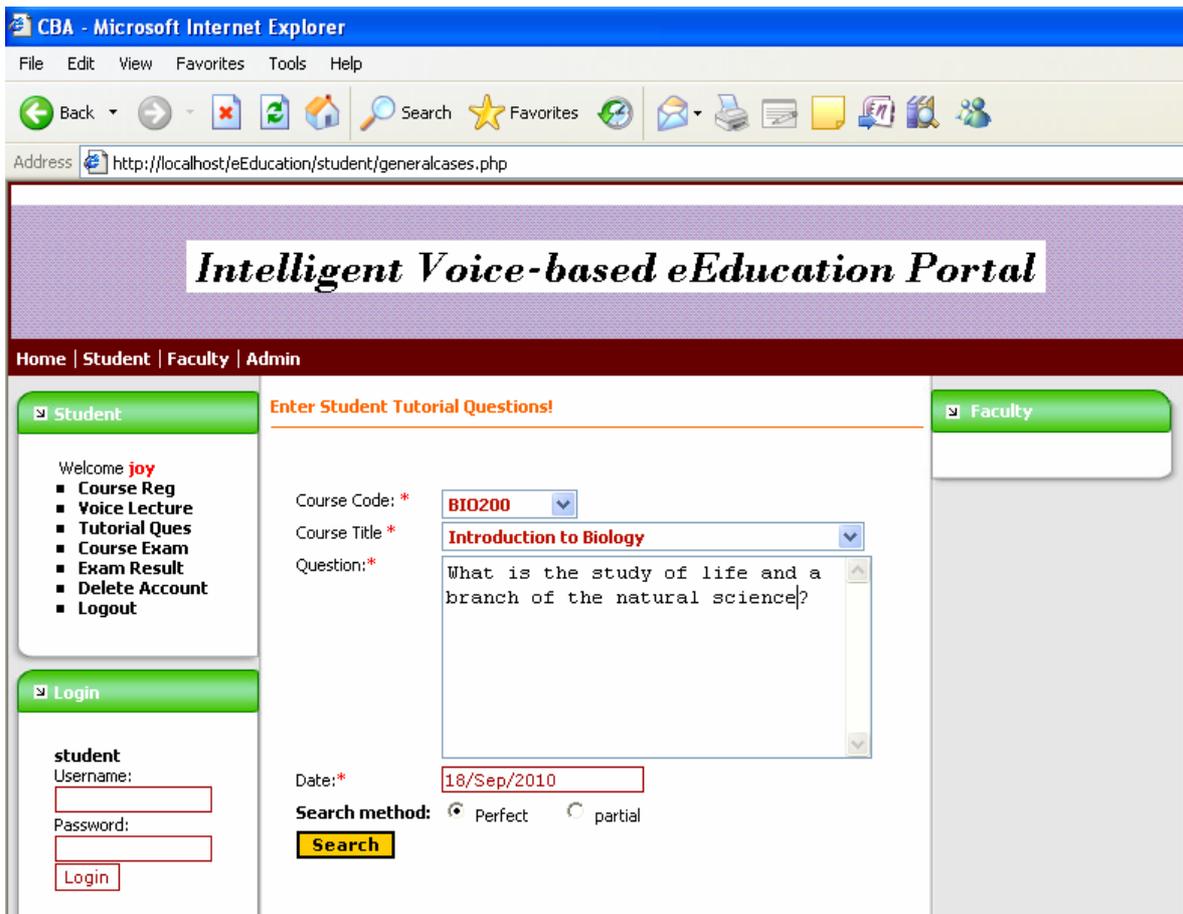
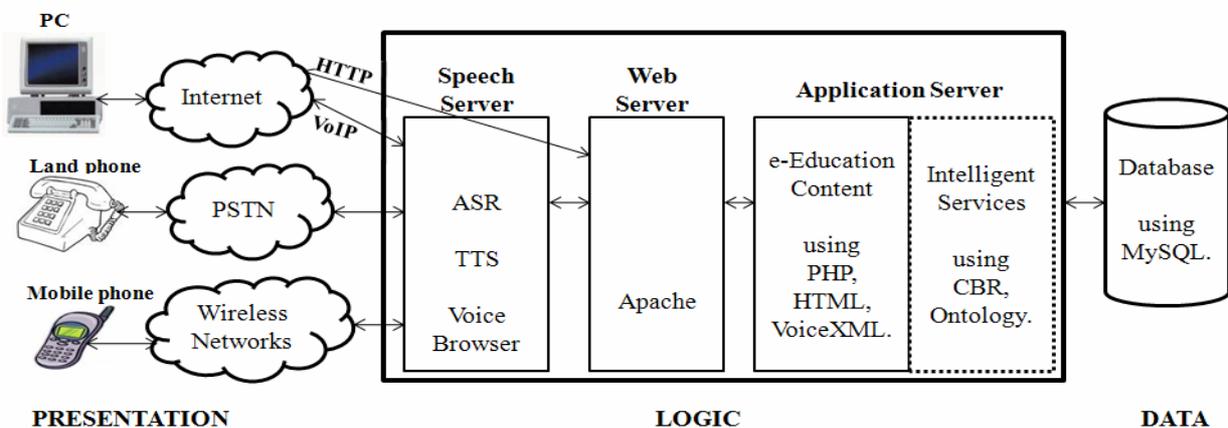**Fig. 3 – A WUI screen shot for tutorial session**

Fig. 4 – Deployment architecture for the framework

The voice browser simply receives any call into the application and submits them to the speech server for further processing. The logic tier comprises of speech server, web server and application server. The broken square box in the application server represents the main contribution of this study. Once a user has been authenticated, the user's query is translated by the automated speech recognition (ASR) to text and passed to the web server for execution. The text-to-speech (TTS) does the reverse of translating text to speech. The client application interfaces with the logic tier using the ASR and TTS. The data tier provides data services and is responsible for changing, adding, or deleting information in the database within the system. The VoiceXML application was deployed on a Voxeo voice server [18] on the web and accessed through VUI from mobile or land phone using the format: <source country int. dial out #> <destination country code><destination area code><generated Voxeo voice network 7 digit #>. To connect to the application, dial: 009-1-312-3805870 or 009-1-412-

5284985 from anywhere in the country.

5. SYSTEM EVALUATION

The application was evaluated for product usability to determine the level of *effectiveness*, *efficiency* and users' *satisfaction*. The evaluation of a product is a fundamental requirement in determining the practical usability of a product [19]. The usability of the e-Education application was measured to specify the features and attributes required to make the product usable using International Standard Organization (ISO) standard of usability [20].

The telephony application was tested using user survey to evaluate the usability attributes. The method proposed in [21] called PARAdigm for DIalogue System Evaluation (PARADISE) could also be use to evaluate the system. The choice for ISO's usability criteria is due to our prior experience of respondents in mobile phone usage and to compare the findings of this study with that obtained from the pilot implementation of *iVEES* project [17]. The users who evaluated the application are not novice users; they garnered some experience of VUI application during

the pilot implementation of the project.

An overall score of all the learners was computed for each of the usability dimension by averaging all the ratings on the questionnaire that was used. With the assistance of some of the non visually impaired teachers, the respondents were taken through a short training on how to dial a telephone number from a mobile phone that will connect the learners to the application and how to navigate within the application. The average (AVG) ratings, standard deviation (SD) and variance (VAR) to determine the mean and dispersion of data collected for the usability attributes are presented in Table 1.

Several usability studies suggest that the system with "Excellent Usability" would have 5 as mean rating, "4 as Good Usability", 3 as Average Usability, "2 as Bad Usability" and "1 as Very Bad Usability". It was proposed in [22] that "Good Usability" should have a mean rating of 4 on a 1-5 scale and 5.6 on a 1-7 scale. Therefore, we can conclude that the prototype application developed for the school has "Good Usability" based on the average (AVG) total rating of 4.16.

Table 1. User survey result N=63

Survey Questions	AVG	SD	VAR
Effectiveness			
I was able to complete my task successfully and correctly using the application	3.94	0.90	0.80
The system did not show error message(s) while using it.	4.02	0.85	0.73
I was able to recover from my mistakes easily.	4.16	0.83	0.68
I feel comfortable using the application.	4.21	0.83	0.68
The voice response I got from the system was clear	4.00	0.78	0.61
Average Effectiveness	**4.07**	**0.84**	**0.70**
Efficiency			
Using the system saves me time of learning	4.08	0.81	0.65
I was able to complete my task on time	4.21	0.72	0.52
I didn't have to carry out too many/difficult steps before completing my task	4.35	0.74	0.55
I was well able to navigate the voice user interface on time when using the system	4.41	0.73	0.54
The system's adaptive features were useful	4.35	0.74	0.55
The system's recommendation features were useful	4.38	0.77	0.59
Average Efficiency	**4.29**	**0.75**	**0.57**
User Satisfaction			
The system was easy to learn	4.05	0.83	0.69
The system was easy to use and user friendly	4.13	0.75	0.56
I am satisfied using the system	4.11	0.84	0.71
I feel the system met my need	4.21	0.79	0.62
am satisfied with the performance of the system in accomplishing my tasks	4.13	0.79	0.63
Average User Satisfaction	**4.13**	**0.79**	**0.63**
Resultant Average Rating	*4.16*	*0.80*	*0.64*

6. CONCLUSION

In this paper, a framework for intelligent voice-based e-Education system has been provided. The framework was used as a generic work guideline to develop a prototype intelligent voice-based e-

Education application for the *iVEES* project. The prototype application was tested in a school for the blind, and the result of evaluation presented. The findings show that the users are enthusiastic about using a voice-based telephone learning as another

form of assistive technology to compliment the conventional learning methods for the visually impaired.

The framework would serve as a reference model for implementing telephone-based e-Education applications for normal and visually impaired learners. The application will also assist people with physical access difficulties (e.g. repetitive strain injury, arthritis, high spinal injury) that make writing difficult. It can also be effective for students with reading or spelling difficulties (e.g. dyslexia).

The future research direction, for this paper is two-fold. First, evaluation of the system using PARADISE, and second, inclusion of voice biometrics features as security mechanism to authenticate candidates, particularly for examination.

7. REFERENCES

[1] M. Kastelic and T. Lončarič. A model of introducing e-learning system at vocational college for business secretaries. *Issues in Informing Science and Information Technology*. Volume 4, 2007, pp. 175-187.

[2] R. Oboko, P. W. Wagacha and E. Omwenga. Adaptive Delivery of an Object Oriented Course in a Web-Based Learning Environment. 3rd *International Conference on ICT for Development, Education and Training,* eLearning Africa 2008 Accra International Conference Centre (AICC), Accra, Ghana, May 28th – 30th, 2008, pp. 274-277.

[3] A. Gil and F. J. García-Peñalvo. Learner Course Recommendation in e-Learning Based on Swarm Intelligence. *Journal of Universal Computer Science*, vol. 14, no. 16 (2008), pp. 2737-2755.

[4] M. B. Raghuraman. Design and implementation of V-HELP system – a voice-enabled web application for the visually impaired. A Master's project presented to the faculty of the graduate college in the University of Nebraska in partial fulfillment of requirements for the degree of Master of Science. Major: Computer Science. August 2004.

[5] A. Holzinger, J. Smolle and G. Reibnegger. Learning Objects (LO): an Object Oriented Approach to Manage e-Learning Content. in: Lazakidou, A. (Ed.) Encyclopedia of Informatics in Healthcare & Biomedicine. Hershey (PA): Idea Group Fererence, pp. 89-98.

[6] B. Pakucs. SesaME: A Framework for Personalised and Adaptive Speech Interfaces. *In Proceedings of EACL-03 Workshop on Dialogue Systems*: Interaction, Adaptation and Styles of Management, Budapest, Hungary (2003) 95–102.

[7] M. S. Abdel-Badeeh and A. Rania. A Case Based Expert System for Supporting Diagnosis of Heart Diseases", *AIML Journal*, 2005, Vol. 5, Issue 1, March'05.

[8] P. Gallivan, Q. Hong, L. Jordan, E. Li, G. Mathew Y. Mulyani, P. Visokey and C. Tappert. VoiceXML Absentee System. *Proceedings of MASPLAS'02*. The Mid-Atlantic Student Workshop on Programming Languages and Systems Pace University.

[9] TellMeVoicePortal(2002) http://www.tellme.com

[10] L. F. Motiwalla and J. Qin. Enhancing mobile learning using speech recognition technologies: a case study. *Eighth World Congress on the Management of eBusiness, 2007*. (WCMeB 2007). 11-13 July 2007 pp.18-18.

[11] L. Motiwalla. Mobile learning: a framework and evaluation. *Journal of Computers & E*ducation, 2006.

[12] P. Walsh and J. Meade. Speech enabled e-learning for adult literacy tutoring. *Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies* (ICALT'03).

[13] G. Salton, Wong and C. S. Yang. A Vector Space Model for Automatic Indexing. *Information Retrieval and Language Processing*, 1975, pp. 613-620.

[14] C. A. Kumar, A. Gupta, M. Batool and S. Trehan. Latent semantic indexing-based intelligent information retrieval system for digital libraries. *Journal of Computing and Information Technology*, CIT 14, 2006, 3, pp. 191-196.

[15] R. Weber, K. D. Ashley and S. B. Brüninghaus. Textual case-based reasoning. *The Knowledge Engineering Review*, 2005, Vol. 00:0, 1–00., Cambridge University Press, Printed in UK.

[16] A. A. Azeta, C. K. Ayo, A. A. Atayero and N. A. Ikhu-Omoregbe. A case-based reasoning approach for speech-enabled e-learning system. *2nd IEEE International Conference on Adaptive Science & Technology (ICAST).* December 14-16, 2009, Accra Ghana. Available online at IEEE Xplore database tab 4: http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5375737

[17] A. A. Azeta, C. K. Ayo, A. A. Atayero and N. A. Ikhu-Omoregbe. Application of VoiceXML in e-learning systems. Cases on Successful E-Learning Practices in the Developed and Developing World: Methods for the Global Information Economy. Published by

IGI Global. 2009. Available online at: http://www.igi-global.com/reference/details.asp?ID=35211&v=tableOfContents

[18] Voxeo (2003), Voice voice server, available online at: http://community.voxeo.com,

[19] N. A. Ikhu-Omoregbe. A formal framework for usable operations support systems in e-health. Unpublished doctoral dissertation, 2007, Covenant University.

[20] ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. 1998.

[21] M. Walker, D. Litman, C. Kamm and A. Abella. PARADISE: A framework for evaluating spoken dialogue agents. *In Proceedings of the 35th Annual Meeting of the Association of Computational Linguistics*. ACL 97, 1997.

[22] J. Sauro and E. Kindlund. A method to standardize usability metrics into a single score. ACM, CHI, April 2-7, Portland, Oregon, USA, 2005.

Ambrose Azeta *is a Ph.D student in the Department of Computer and Information Sciences, Covenant University, Ota, Nigeria. He holds B.Sc. and M.Sc. in Computer Science from University of Benin and Lagos respectively. His current research interests are in the following areas:* Software Engineering, Algorithm Design and Mobile Computing. He currently lectures at Covenant University. He is a member of the Nigerian Computer Society (NCS) and Computer Professional Registration Council of Nigeria (CPN).

Charles K. Ayo *holds a B.Sc. M.Sc. and Ph.D in Computer Science. He is currently an Associate Professor of Computer Science and the Director, Academic Planning Unit of Covenant University. He was the immediate past Head of Computer and Information Sciences Department of the University. Ota, Ogun state, Nigeria, Africa. His research* interests include: mobile computing, Internet programming, eBusiness and government and Software Engineering. He is a member of the Nigerian Computer Society (NCS), and Computer Professional Registration Council of Nigeria (CPN). Dr. Ayo is a member of a number of international research bodies such as the Centre for Business Information, Organization and Process Management (BIOPoM), University of Westminster; the Review Committee of the European Conference on E-Government; the Editorial Board, and Journal of Information and communication Technology for Human Development (IJICTHD) amongst others.

Dr. Aderemi Aaron Anthony Atayero, *holds a B.Sc. degree (summa cum laude) in Radio Engineering and M.Sc. in Satellite Communication Systems (1992, 1994 respectively) both from the Moscow Institute of Technology, and a Ph.D. in Technical Sciences (Speech* Processing / Satcom, 2000) from the Moscow State Technical University of Civil Aviation, Moscow, Russia. His current area of research interest is as follows: Speech Processing. FPGA Design, Architecture, and Applications: FPGA – Implementation of Digital Speech Processors. FPGA – Implementation of Digital Filters. FPGA System–on–Programmable Chip Design. He is a member of a number of academic and professional organisations including the Institute of Electrical and Electronic Engineers (IEEE), Nigerian Association of Inventors (NAI). Dr. Atayero is a Senior Lecturer in the department of Electrical and Information Engineering at Covenant University.

Nicholas Ikhu-Omoregbe, *has a B.Sc degree in Computer science from the University of Benin, Benin city, an M.Sc degree in Computer Sciences from the University of Lagos, and a PhD degree in Computer Science from Covenant* University, Ota, Nigeria. His research interests include: Software Engineering, Mobile Computing, Mobile Healthcare and Telemedicine Systems, and Soft Computing. He currently lectures at Covenant University. He is a member of the Institution of Electrical and Electronics Engineers (IEEE).