# Chapter 4
# A Voice-Enabled Framework for Recommender and Adaptation Systems in E-Learning

**A. A. Azeta**
*Covenant University, Nigeria*

**C. K. Ayo**
*Covenant University, Nigeria*

**N. A. Ikhu-Omoregbe**
*Covenant University, Nigeria*

## ABSTRACT

*With the proliferation of learning resources on the Web, finding suitable content (using telephone) has become a rigorous task for voice-based online learners to achieve better performance. The problem with Finding Content Suitability (FCS) with voice E-Learning applications is more complex when the sight-impaired learner is involved. Existing voice-enabled applications in the domain of E-Learning lack the attributes of adaptive and reusable learning objects to be able to address the FCS problem. This study provides a Voice-enabled Framework for Recommender and Adaptation (VeFRA) Systems in E-learning and an implementation of a system based on the framework with dual user interfaces – voice and Web. A usability study was carried out in a visually impaired and non-visually impaired school using the International Standard Organization's (ISO) 9241-11 specification to determine the level of effectiveness, efficiency and user satisfaction. The result of the usability evaluation reveals that the prototype application developed for the school has "Good Usability" rating of 4.13 out of 5 scale. This shows that the application will not only complement existing mobile and Web-based learning systems, but will be of immense benefit to users, based on the system's capacity for taking autonomous decisions that are capable of adapting to the needs of both visually impaired and non-visually impaired learners.*

DOI: 10.4018/978-1-4666-2208-1.ch004

## INTRODUCTION

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 places strong emphasis on learners' previous knowledge. The uniqueness of a learner is met by making the Web-based learning content adaptive (Oboko et al., 2008). 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 interest is cumbersome (Gil and García-Penalvo, 2008). The process of selection may be easier with the normal users, but for certain category of learners with a visual impairment, 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 VUI applications is Voice eXtensible Mark-up Language (VoiceXML or VXML).

VUI applications are primarily developed to cater for the visually impaired (Raghuraman, 2004), to address the problem of Web accessibility associated with the use of m-Learning. The major problem of Web access is that the services of reusable learning objects currently available to Internet-connected users are not available for the visually impaired (Holzinger et al., 2006), for reasons of constraint in their navigation abilities. Thus, the voice-based Web applications mainly designed for the visually impaired lack adequate quality of using recommender and adaptable system for learning, which is a major requirement for this category of users as a result of their visual impairment. Recommender systems are software agents that recommend options for users. These agents can be very useful in an E-Learning environment to recommend actions, resources or links (Zaiane, 2005). The services provided by recommender systems are generally referred to as recommendation services. Adaptive Web system monitors particular user's behavior and characteristics. Based on them, the system compiles document from larger universal source document and then produce adapted document to users (Bures & Jelinek, 2004).

The goal of adaptive voice-based Web learning is to adjust the content of the learning objects to suit user's knowledge level, whereas recommendation services provide the most appropriate learning objects to users through voice. The inability of existing E-Learning 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-Learning systems, to make them more adaptable and provide recommendation services to users' needs. Thus, we have to further take into account that the reference guidelines for developing existing voice-based E-Learning applications, which although important, lack intelligent component services to approach the problem. One of the ways of enhancing existing voice-based E-Learning applications is to adapt their content to the needs of each student.

The needs of users differ when accessing learning content, such needs may require the ability of the system to reason, make decisions, be flexible and adapt to divers requests during interaction. These needs have placed new requirements in voice application development such as the use of advanced models, techniques and methodologies which take into account the needs of different users and environments. The ability of a system to behave close to human reasoning for adaptability support is often mentioned as one of the major requirements for the development of intelligent voice applications. Intelligent systems belong to the domain of Artificial Intelligence (AI). AI is a branch of computer science concerned with creating or mimicking intelligent behaviours or thoughts in computers (Gloss, 2004). A system can also be said to be intelligent if it is able to improve its performance or maintain an acceptable level of performance in the presence of uncertainty. The main attributes of intelligence are learning, adaptation, fault tolerance and self-organisation (Jain and Martin, 1998). If adaptability services must exist as a feature of voice-based E-Learning system, then an intelligent agent will be required to power the application.

From the foregoing, it becomes obvious that there exists a need to provide a framework for recommender and adaptation system in E-learning. This study aims at addressing the needs as identified, and employing the framework so obtained for the development of a prototype voice-enabled framework for recommender and adaptation system *(VeFRA)*. The prototype application developed was tested in a school for the blind, and the result of evaluation reported. The framework suffices as a reference model for implementing intelligent voice applications in the domain of E-Learning. Applications developed based on this framework would exhibit the necessary attributes of Case Based Reasoning (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' requests based on requirements. The application will therefore be helpful for people with physical access difficulties engendered by their visual impairment.

The objective of this study is to provide a Voice-enabled Framework for Recommender and Adaptation (*VeFRA*) Systems and adapt the framework for the development of *VeFRA* that has the capability to improve learning processes using telephone and Web-based technologies.

## BACKGROUND

This section discusses existing frameworks and systems in the area of E-Learning. A framework provides reference guideline for developing applications. Existing framework for developing voice-based E-Learning system such as VoiceXML application development life-circle (VoiceXML, 2007) does not have an intelligent component services and it was modeled after the conventional software development life-circle model. However, VoiceXML applications developed using such framework serve as an alternative platform for non-visually impaired users.

## REVIEW OF EXISTING SYSTEMS

In addition to the provision of alternative platform for normal users, voice-enabled E-Learning systems can be helpful for people with physical access difficulties (e.g. repetitive strain injury, arthritis, high spinal injury) that make writing difficult (Donegan, 2000). It can also be effective for students with reading, writing or spelling difficulties, such as dyslexia (Husni & Jamaludin, 2009) and for those with visual impairment (Nisbet & Wilson, 2002). Paul (2003) described some important factors to be considered when introducing and using speech recognition for students with disabilities. It was reported by Paul (2003) that speech recognition programs have been used successfully by students with a range of disabilities. The programs have been shown to provide an effective means of writing and recording work, and in some cases, they produced significant improvements in basic reading, spelling and writing skills. Chin et al.,

(2006) recommended that one can actually make use of VoiceXML technology to build speech applications that can serve educational purposes or in other words, to build an online learning system that provide better accessibility to users. One of the E-Learning applications that can be provided using speech technology are those that deliver basic teaching by simply listening. For example, students can check their scores or other information by simply calling a particular number and get the information they want. Several schemes for providing E-Learning system have been proposed in literature. Chin et al., (2006) went further to develop an E-Learning prototype based on VoiceXML concepts. However, the prototype was not deployed on the real telephone environment. Instead, a microphone and a speaker were used to simulate the telephone environment.

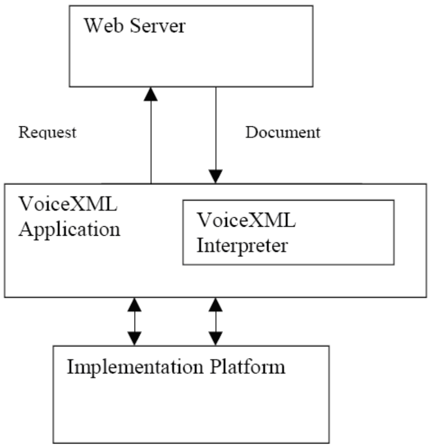
Gallivan et al. (2002) developed a voice-enabled Web-based absentee system on TellMe voice portal (TellMe Voice Portal, 2002). It was tested by the class of software engineering students; students who intended to miss a class called the VoiceXML telephone number and were led through an automated dialog to record their names, the date and time of the call, the courseID, and the date they would miss class in a database. The system provides the instructor and other administrators with a permanent record of absentees that can be accessed and displayed in various forms through a Web interface. The absentee system application was developed basically for Pace University students to report class absences. The VoiceXML absentee system has been designed to include record keeping of absentee calls from students, faculty and university staff. This system made use of the general VoiceXML platform architecture in Figure 1.

Likewise, Raghuraman (2004) designed, implemented and deployed a voice-enabled application called V-HELP system. In the V-HELP system, a portion of the Computer Science and Engineering (CSE) department website was voice enabled using VoiceXML to enable the visually impaired student population to access the information on it. The website also provides easy access to the information through voice, while the user is on the go, and if the user does not have a handheld with sufficient screen size or is constrained by the bandwidth to browse through the visual content. It can be used by visually impaired student population of the school to access information about the Computer Science and Engineering department. Though the primary purpose of this project is to utilize information technology in the area of assistive technology, it can also serve students, who do not have Internet connection to browse the text based website or who do not have access to a computer. Since this Web portal can be accessed through a phone call, students will also be able to access CSE department information while they are on the go.

Similarly, Kolias et al., (2004) described the design and implementation of an audio wiki application accessible via the Public Switched Telephone Network (PSTN) and the Internet for educational purposes. The application exploits World Wide Web Consortium standards such as VoiceXML, Speech Synthesis Markup Language (SSML) and Speech Recognition Grammar Specification (SRGS). The purpose of the application was to assist visually impaired, technologically uneducated, and underprivileged people in accessing information originally intended to be accessed visually via a Personal Computer. Users access wiki content via wired or mobile phones, or via a Personal Computer using a Web Browser or a Voice over IP service. In its current form, the application supports pronunciation of the content of existing articles via synthetic speech. It does not support editing or creation of new articles due to the incapability of existing recognition engines to provide full speech recognition instead of recognition of words from a predefined set.

Motiwalla and Qin (2007) explored the integration of speech or voice recognition technologies into m-Learning applications to reduce access barriers. An educational online forum accessible

*Figure 1. Architecture for voice-enabled web-based absentee system (source: Gallivan et al., 2002)*




through mobile devices was built, based on an m-Learning framework proposed in Motiwalla and Qin's previous work. A customized Interactive Voice Response (IVR) and Text-To-Speech (TTS) technologies to allow users to interact with the forum through voice commands was also developed. This voice-enabled discussion forum application does not only help normal users to avoid the cumbersome task of typing using small keypads, but also enables people with visual and mobility disabilities to engage in online educations. The prototype forum was tested in two different blind institutions in Massachusetts with 10 users. The results from this study provide insights into how to improve accessibility of m-Learning. The research utilizes the framework in Figure 2 to develop a set of mobile phone accessible applications that can be used for the m-Learning environment.

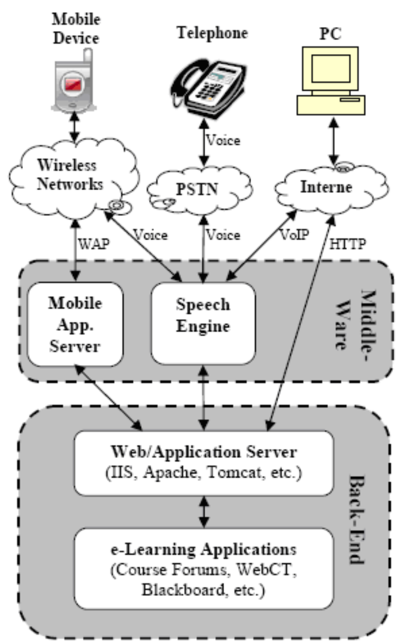
The architecture in Figure 3 utilises the m-Learning framework in Figure 2. The architecture contains three layers: back-end layer, middle-ware layer, and user interface layer. The back-end layer consists of core Web applications such as forums, Web blogs, Wikis, as well as E-Learning applications such as WebCT, Blackboard, etc. The back-end also consists of the Web/application server that supports those applications. The back-end provides source materials and basic services of the whole m-Learning system. The middle-ware layer consists of components that enable the users to access back-end E-Learning applications through various means. For example, the speech engine may contain customized IVR and TTS technologies or commercial speech recognition software that allow users to interact with the E-Learning applications through voice commands. The mobile application server component enables users to interact with the E-Learning applications through WAP or SMS on their mobile devices. Finally, the user interface layer resides on user devices. Users can access the E-Learning applications on their PCs, telephones, or mobile devices. They can do so by typing messages in text browsers or issuing voice commands.

Also, Azeta et al. (2008c) carried out a preliminary study of the *VeFRA* project on the design and development of a prototype telephone-based E-Learning portal for course registration and

*Figure 2. An m-learning framework (source: Motiwalla, 2007)*

| | Personalized Content | Collaborative Content | |
|---|---|---|---|
| **PUSH Mechanism** | *Pedagogical Agents & Mentors* | *Communication Aids* | *SMS, IM, Alerts, Scheduling Calendars* |
| **PULL Mechanism** | *System Tools & Resources* | *Simulated Classrooms* | *WML websites, Discussion Boards & Chat Forums* |
| | *Alerts, Scheduling Calendars, WML websites* | *SMS, IM, Discussion Boards & Chat Forums* | **M-learning Applications** |

*Figure 3. M-learning architecture with speech recognition enhancement (source: Motiwalla & Qin, 2007)*




examination by the students. Jianhua et al., (2009) employed the technology of VoiceXML to construct an open VoiceXML-based Mobile Learning System (VMLS). The hardware and software system structure design was achieved. The research results could not only be used in VMLS, but could also be used in other voice-based application system. Peter et al., (2008) presented an architecture to enable multi-channel interaction (Web, SMS/MMS and voice) with services that support collaboration on a blog server. The use of blogs to support learning can be found in some works by Lin et al. (2006) and Divitini et al. (2005). The motivation for this work was to enable learners to access collaborative learning services regardless of their situation and location. Kondratova (2009) discusses issues associated with improving usability of user interactions with mobile devices in mobile learning applications. The focus is on using speech recognition and multimodal interaction in order to improve usability of data entry and information management for mobile learners.

A Voice-enabled Interactive Service (VoIS) was proposed by Motiwalla (2009) for E-Learning. The goal of VoIS platform for an E-Learning system is to better meet the needs of students and mobile learning communities via speech recognition technology. The pilot study explored this goal by demonstrating how interaction tools like discussion forums can be voice-enabled using VoiceXML. The authors design, develop, and evaluate a prototype application with blind and visually impaired learners from different organizations.

The study by Garcia et al., (2010) introduced ‘‘Voice Interactive Classroom,’’ a software solution that proposes a middleware approach to provide visual and auditory access to Web-based learning. The approach has successfully enabled the integration of visual interfaces to voice dialogues and to provide auditory access to functionalities already present in various learning management systems such as Segui, Moodle and Sakai.

The educational contents and the framework in these systems lack support for intelligence such as the ability to take decisions that will adapt to users' request. Thus, we believe that if finding suitable content of learning resources can be introduced into a voice-based E-Learning system, including CBR and ontology-based properties, then the voice-based E-Learning system will have the capabilities of providing adaptable and intelligent learning to users.

This study also addresses some of the concerns expressed by Devedzic (2004), where it was noted that there are several challenges in improving Web-based education, such as providing for more intelligence, and that a key enabler of this improvement will be the provision of ontological support. Other studies such as Holohan et al., (2005); Aroyo and Dicheva, (2004); Aroyo et al., (2003) have used ontology to provide E-Learning

Content Generation based on Semantic Web Technology. Meanwhile these approaches were not based on voice architecture.

## MOTIVATION FOR THE STUDY

The equipment used as assistive technology for learning in school for the blind like Slate and Stylus, Mathematics board and Figures, Braille, Typewriters, Abacus, and so on are expensive to procure and maintain (Azeta et al., 2009b). As a result of the pace of technological revolution, these pieces of equipment are gradually being replaced by the numerous applications developed to run on the Internet.

The Internet, certainly has received significant attention in recent years, but verbal communication is still the most natural form of interaction amongst human. Therefore the human speech is a very good medium of communication using telephone access. Given the naturalness and expressive power of speech, speech input and output have the potential for becoming key modalities in future interactive systems (Niels, Ole, & Bernsen, 2000). There is an effort to use this kind of communication also in human–computer interaction and for data access on the Web (Ondas, 2006). Mobile phone is one of the media for accessing Web content. The limitations of mobile phones such as small screen size, small keypad, and limited battery power to effectively hold and display information does not pose any constraint to telephone calls and conversation, compared to when using it to access the Web through Wireless Application Protocol (WAP). Typing on the tiny keyboards of WAP-enabled phones or pointing with stylus is very uncomfortable and prone to errors. Another issue is that mobile devices are often used when a person is really ”on the move”. Operating in such conditions is either impeded or even prohibited, e.g. in the case of car driving (Zaykovskiy, 2006).

The natural way to overcome this problem is the use of voice technology. The possibility of using spoken language to interact with computers has generated wide interest in spoken language technology. Many universities and research laboratories have put their efforts to a development of the spoken language technology (Lerlerdthaiyanupap, 2008). However, voice technology is limited to only one form of input and output – human voice (Kondratova, 2009).

The high popularity and wide spread use of telephone globally, with an estimate of about four billion active telephone lines (fixed and mobile phone) prompted the World Wide Web Consortium (W3C) speech interface framework (Larson, 2000) to create an avenue for users to interact with Web-based services via key pads, spoken commands, prerecorded speeches, synthetic speeches and music. The interface framework led to the massive convergence of Web and voice technologies to provide solutions in the areas of E-Health, E-Stock, E-Democracy, E-Learning and a lot more.

Majority of the present day E-Learning applications have support for the Web user interface (WUI) through the use of PC, while others are based on mobile devices such as Wireless Application Protocol (WAP) phones and Personal Digital Assistants (PDAs). Given the limited input/output capabilities of mobile devices, speech presents an excellent way to enter and retrieve information either alone or in combination with other access modes. Furthermore, people with disabilities should be provided with a wide range of alternative interaction modalities other than the traditional screen-mouse based desktop computing devices. People who suffer disability whether temporary or permanent, either in reading difficulty, visual impairment, or any difficulty using a keyboard, or mouse can rely on speech as an alternate approach for information access (Farinazzo et al., 2010).

The need to reduce access barriers for the visually impaired has necessitated this research in the education domain with mobile and land phones as access point. The problems with the various conventional learning systems such as face to face, telephone, electronic mail, chat room, instant messaging, etc, are numerous, since using them becomes a more difficult task for those with disabilities.

Lack of accessibility in the design of E-Learning courses continues to hinder students with vision impairment. Despite significant advances in assistive technologies, blind and visually impaired Internet users continue to encounter barriers when accessing Web content (Gerber & Kirchner, 2001; Williamson et al., 2001). E-Learning materials are predominantly vision-centric, incorporating images, animation, and interactive media, and as a result students with low or acute vision impairment do not have equal opportunity to gain tertiary qualifications or skills (Armstrong, 2009). A blind person cannot see or communicate through mail or electronic means that require the ability to see the screen. Lack of provision for voice in the conventional learning methods has restricted support for people with limited capabilities such as the visually and mobility impaired that affect either data entry, or ability to read (thus check) what they have entered, since these applications are visual in nature and require the ability sight to see the blackboard or computer screen and manipulate the computer keyboard. More so, it excludes people who have medical history of Repetitive Strain Injury (RSI) as a result of their sedentary lifestyle using a peripheral device such as a keyboard. The Web-based and mobile learning system does not have support for people with textual interpretive problems (e.g. dyslexia) to enter text verbally, and others who have problems of identification of characters and words.

An obvious limitation in providing Information Technology access to learners with low vision is the difficulty that these learners have in reading text. This difficulty is compounded when a visually impaired user is faced with computer technologies and concepts, as a first time user. Voice technology is a possible solution to this problem as it can enable students to learn literacy skills using a more natural human-computer interface (Walsh & Meade, 2003). However, existing voice-based learning application have not been able to adequately proffer solution to the problem of providing suitable learning content in the midst of numerous learning resources available on the Web for both normal and visually impaired learners.

## STATEMENT OF THE PROBLEM

In some Educational Institutions, particularly public schools in the developing nations, there is growing shortage of academic staff and the demand for teachers has continued to increase. Demand for access to education remains much higher than the system's physical capacity to accommodate students in public schools (NUC, 2004). As a result, the ratio of teachers to students is constantly unbalanced with students having an increasing number. The process involved in conducting course registration, lecture delivery, examination and result is tedious, since many students are attended to by few teachers. More so, no adequate attention is given to students by their teachers in the areas of counseling, teaching, invigilation, etc. Currently, several Universities and Colleges are overcrowded with inadequate facilities to match the upsurge, thus resulting in deviation from the approved carrying capacity accomplished with poor academic performance (Bobbie et al., 2008).

While trying to address the aforementioned problems, many studies have shown the value of existing platforms such as Web and m-Learning, but lack access to such applications for the large number of students through virtual learning. The number of E-Learning Web portal has continued to rise as a result of the advancement in Web technologies. Some of the popular Web portal available for learning includes: Moodle, WebCT, Blackboard and Claroline (Ermalai et al., 2009). Voice-based learning (Motiwalla & Qin, 2007) evolved to address the problems of accessibility associated with m-Learning, arising from the limitation in using mobile phone to access learning content on the Web.

With the proliferation of Web learning resources, finding suitable content (using telephone) has become a rigorous task for voice-based online learners to achieve better performance. The biggest disadvantage of voice-based technologies is the rigid structure that they impose on the end user. While it is convenient to use mobile telephony application, it can be extremely slow when the user is forced to drill through several layers of options before finding exactly what he/she wants (Raghuraman, 2004). The Finding Content Suitability (FCS) problem with voice E-Learning applications is more complex when the visually impaired learner is involved. Thus, existing voice-enabled E-Learning applications lack the necessary attributes of reasoning and adaptability during interaction to be able to address the FCS problem.

To solve the FCS problem and enhance voice-based learning, recommending and adapting the contents to the needs of each student is required. It is a necessity that spoken dialogue system performance be enhanced through adaptation (Pakucs, 2003). More so, the adoption of artificial intelligence methodologies using CBR and domain ontology to provide recommendation services to the different needs of students is essential (Abdel-Badeeh and Rania, 2005). Generally, there is dearth of voice application in the domain of E-Learning and the few existing voice-based E-Learning application are not intelligent to provide recommendation services beyond student's queries and request.

Thus, we take cognizance of the fact that there is no reference model containing intelligence upon which the existing voice-based E-Learning applications were developed. Although the approach employed for developing some of them are the general software development life-cycle model, they do not constitute a framework to provide intelligence for voice E-Learning applications.

## THE VEFRA FRAMEWORK

This section presents requirement elicitation and *VeFRA* framework. The components of *VeFRA* framework including adaptation and recommendation services are discussed. An Education Ontology was designed using Protégé 3.4.1 and developed with Web Ontology Language (OWL).

## ELICITATION OF REQUIREMENTS

The development of a framework requires domain information from the potential users of the system. A school for visually impaired and non-visually impaired were visited to enabled us acquire the required information. The investigation carried out in the visually impaired school shows that the conventional assistive equipment used are expensive to acquire and maintain. It is also difficult to organize the visually impaired learners into classes from their hostel without the assistance of attendants, and this make the entire learning process to be cumbersome.

From investigation carried out in schools attended by the normal learners (i.e. students from non-visually impaired institutions), the existing learning methods in most public institutions in Nigeria are yet to sufficiently provide E-Learning access to students to enable them learn independently (i.e. on their own or on the move) irrespective of location. Students have to be physically present in the classroom to participate in class lectures. This does not favor most of the working class students who have to constantly leave the office for lectures, tutorial and examination. It is the opinion of students and management of one of the campuses/centers that a system that allows the students to learn on their own through distance learning would minimize the problem of over-enrollment of students and scarcity of teachers, among others in the centers.

The collated requirements and specification for the system include: registration of courses, audio streaming of lecture notes, tutorial questions and answers, course examination and checking of examination results. In order to activate the intelligent component services of the system, discussions were held with some students and teachers involved in the learning process to gather the required information that is relevant for CBR paradigm. The system requirement is presented in Figure 4, in which the outer circle is the actual needs of the learners. The inner circle is the platform upon which the framework was developed and deployed.

## THE VOICE-ENABLED 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-Learning 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 E-Learning application, participate in tutorials and check examination results. As was expected, beginners 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: 1) Ability to provide users with services involving course registration, voice lecture, tutorial, examination and result; 2) ability to accommodate different types of users based on their respective learning profile; 3) ability to differentiate the course contents into different levels such as expert, intermediate and beginners; and 4) selection of different course contents based on users' profile. The proposed framework for *VeFRA* is shown in Figure 5. The framework comprises of interaction and intelligent layer.

Learner's knowledge level is subdivided into three categories: beginner, intermediate and expert. Learner profile determines whether a learner should receive beginner, intermediate or expert content of lecture, tutorial, examination and result module. The system navigation process is contained in the interaction management. 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 to provide recommended answers to tutorial questions using previous experience of learners, and also to expand the search for answers using stemming algorithm. Learning objects consist of chunks of course materials in text format which allow information to be presented in several ways to users. Tutorial questions asked by previous learners and E-Learning data are stored in case knowledge and domain ontology respectively.

*Figure 4. The system requirement for VeFRA framework*

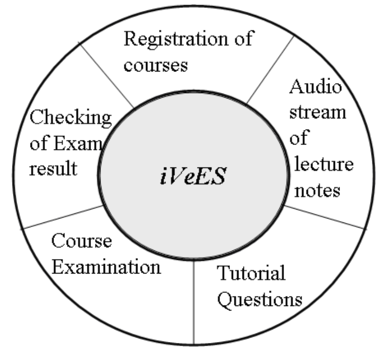
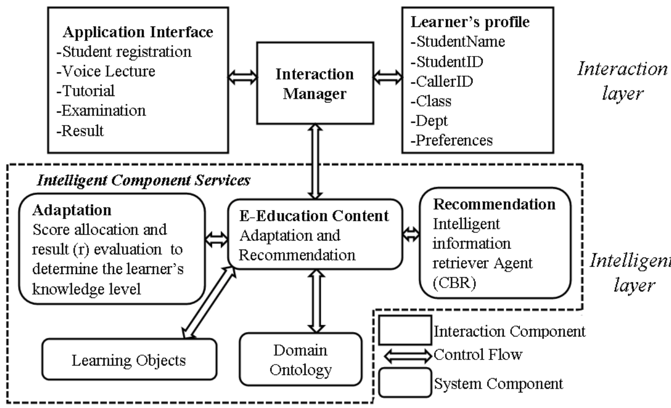


*Figure 5. Voice-enabled framework for recommender and adaptation (VeFRA) systems*




## The Adaptation Component Service

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 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 Figure 6.

Score allocated for each question:

$$Q = \sum_{k=1}^{q} Q_k \quad \text{and} \quad S = \frac{Max(r)}{Q} \tag{4.1}$$

Evaluation result for each Activity:

$$Y = \left( \sum_{j=1}^{s} S_j \right) + W \tag{4.2}$$

Evaluation result for each knowledge level:

$$R = \sum_{m=1}^{y} Y_m \tag{4.3}$$

The following assumptions were made in formulating Equations 3.1, 3.2 and 3.3.

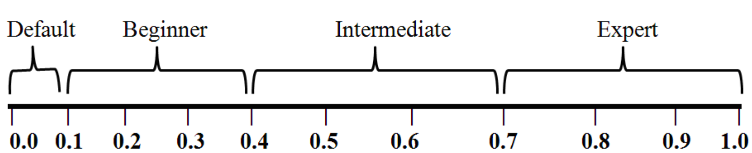
r = Expected result for score evaluation criteria (see Figure 14).

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.

*Figure 6. A model of score evaluation criterion*



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 Component Service

The recommendation service was designed using CBR, stemming and domain ontology. One of the main challenges is to establish structures with which data can be represented as cases. This structure is presented in Figure 7. Alphanumeric data are stored as *text,* numeric data as *number* and date as *date* attribute. A typical data entry for the structure would be 011122, BIO200, "Introduction to Biology", "BiologyCase", "What is Biology", "Biology is the study of life and a branch of the natural sciences", "09-04-2010", "partial".

CBR may be viewed as a cycle involving four steps (Spasic et al., 2005): *retrieve* the most similar case, *reuse* the case to solve the new problem, *revise* the suggested solution and *retain* the useful information obtained during problem solving, after the solution has been successfully adapted to the target problem. The retrieval stage is the first step. The procedure of case retrieval begins with identifying the most important features and using them in identifying cases to reuse. Additional domain information often improves results, i.e. a list of words and their synonyms or a dictionary provides comparable words. Our proposed system uses Ontology for Voice-based Education that represents specific knowledge, i.e., relationship between words used in Education (Biology) subject. The text tokenizer and Modified Porter Stemming (MPS) 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.

A UML sequence diagram for the recommendation service is shown in Figure 8. The learner initiates voice response for tutorial questions and the learners' characteristics are updated. The request for tutorial question is executed using search agent, which then provides the recommended answer using some rules. Thereafter, the recommended answer is passed to the leaner through voice response.

## ONTOLOGY-BASED MODEL OF VEFRA FRAMEWORK

Domain ontology is a detailed description about an application specific domain including concepts, entities, attributes and processes related to a given application domain. In this research, our domain is *Education Ontology*. The proposed ontology consists of E-Learning objects and their relationships. Therefore via these ontological relationships, the required learning resources can be located and their metadata can be retrieved. Then matching of user profiles and learning resource metadata can be done to ascertain the relevance of them. This ontology can be extended according to the changes to user profiles and learning content. When we try to satisfy learner requirements we

*Figure 7. A case structure for VeFRA*

| Case field | Attribute |
|---|---|
| StudentID: | number |
| CourseID: | text |
| CourseTitle : | text |
| CaseName: | text |
| Query: | text |
| Solution: | text |
| Date: | date |
| SearchMethod: | text |

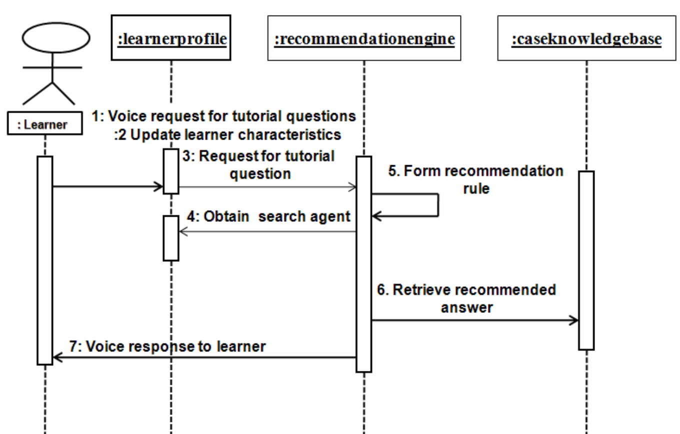
need to match the user profiles with the learning objects in the system (Heiyanthuduwage et al., 2006). However, a direct matching between the learning objects and user profiles is not efficient enough and it will take lot of memory.

## THE EDUCATION ONTOLOGY DEVELOPMENT

Our first goal was to collect and structure the available information related to the use of ontologies in the field of education. In this study, we have to deal with subject domain ontology and not structure ontology. The *domain ontology* represents the basic concepts of the domain under consideration along with their interrelations and basic properties. The *structure ontology* defines the logical structure of the content. The Web Ontology Language (OWL) was used to design our ontology. We created an ontology named Education.owl and three classes that model separately the different components of the system. The ontology is encoded in OWL using Protégé 3.4.1. In the ontology there are classes, relationships and attributes related to the domain. New classes, relationships and attributes can be entered to the ontology e.g. by using Protégé 3.4.1. Each model can be extended to include as many relationships and classes as possible. The knowledge base resides in Apache HTTP Server. Figure 9 shows a snapshot of our Education ontology with the classes and properties. It has three classes; Learner, LearningContent and Assessment. A child node is an instance of a father node i.e. Course and Tutorial are instances of LearningContent. All the instances of Enumeration applies to all the classes. The ontology can be used for adaptive learning to retrieve the content of a course. The architecture of Education Ontology model is presented in Figure 10, with the three classes and the relationships between the classes.

Once a class of the ontology is selected, the application delivers the instances related to that class. Figure 11a and 11b presents Classes in Education.owl Ontology and the properties/rela-

*Figure 8. A UML sequence diagram for recommendation service*



tions respectively using Protégé 3.4.1. The development of the Education ontology involves the creation of the classes, properties and relationship between deferent classes.

## VEFRA APPLICATION ARCHITECTURE

The VoiceXML *VeFRA* architecture (see Figure 19) consists of some VoiceXML software suites including Registration, Voice lecture, Tutorial, Examination and Results. For this prototype implementation, we used free Voxeo hosting service to provide a quick way to get started in order to eliminate huge capital expenses in subscription.

The *VeFRA* application architecture shown in Figure 12 was drawn using modular architectural design. The description of the architecture is presented as follows:

1. The learner connects through a telephone and PSTN to the VoiceXML interpreter through Voxeo Speech gateway.
2. The VoiceXML interpreter execute the call interaction with a caller using the instructions of a VoiceXML script supplied by the voice commands in application server.
3. The interpreter calls TTS and ASR as plugins to complete its tasks.
4. The VoiceXML interpreter communicates via Web protocols (HTTP) to VoiceXML application server.
5. The VoiceXML application server delivers the application including VoiceXML text pages.
6. The Web application server queries the database via apache to dynamically retrieve information and the VoiceXML interpreter TTS speak with the caller.

*Figure 9. Main classes of the education.owl ontology*

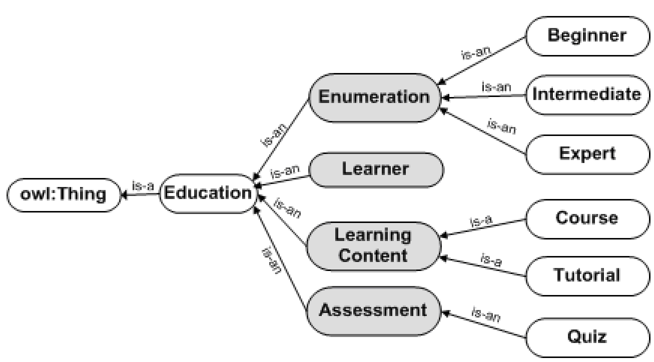



*Figure 10. Architecture of education ontology model*

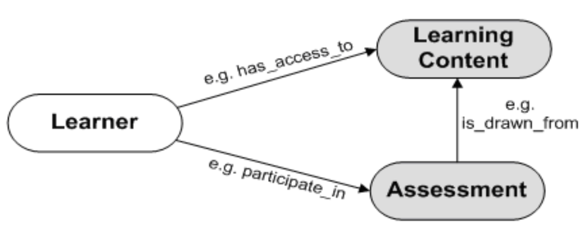
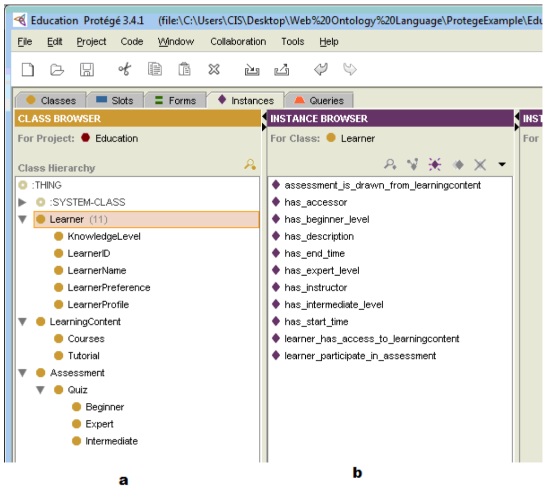


*Figure 11. A screenshot of education ontology design using Protégé 3.4.1*

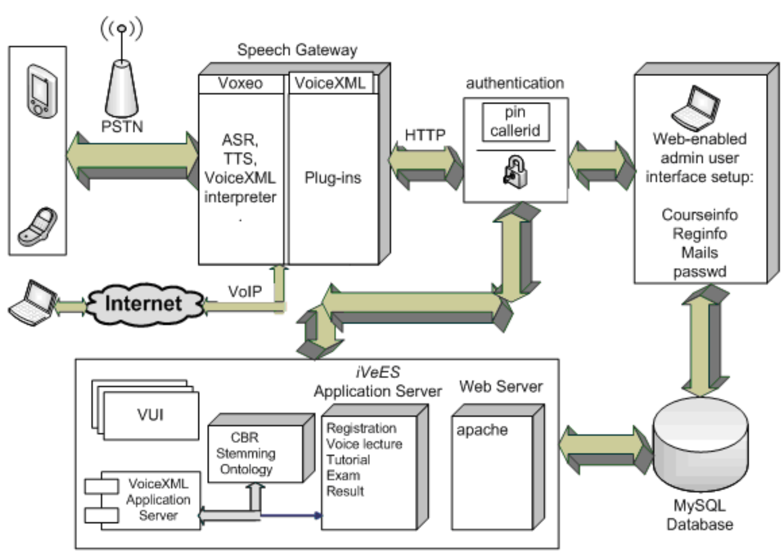



## SYSTEMS IMPLEMENTATION

A prototype application has been developed using the *VeFRA* framework. The components of the application comprising design, implementation and development have been reported in Azeta et al., (2009b) and Azeta et al., (2009c). Figure 13 contains screen shot of a sample tutorial question used to demonstrate the recommendation component service of the application. The faculty or course lecturer select course code, course title, and type tutorial question for which answers are sought. The date is automatically selected by the system. It is the responsibility of the faculty to select the search method as either partial of perfect. Partial means displayed answers will include both the exact match and matches that are similar, whereas perfect search showcased answers that are exact. The WUI is developed for system administration purposes such as upload of tutorial questions and lecture notes. Voxeo speech platform (Voxeo, 2003) was engaged as the speech server while CBR and Modified Porter Stemming (MPS) algorithm were used to provide intelligent services.

A prototype part of the VoiceXML application (voice user interface) was deployed on a Voxeo voice server (Voxeo, 2003) on the Web and accessed from a mobile phone and land phone using the format:<source country int. dial out #> <destination country code><destination area code><generated voice network 7 digit #>. Figure 14 shows a model of mobile phone interaction with the system. A user dial a telephone number in Figure 14a. Once a connection is established

*Figure 12. VeFRA application architecture*



(Figure 14b), the user can then commence conversation (Figure 14c) with *VeFRA*.

Similarly, a prototype part of the Web application (Web user interface) was deployed on a Web server on the Internet and can be accessed using an Internet ready desktop PC by visiting the URL - http://www.emma.byethost24.com/eEducation. After deploying the application, the teachers and students went through some training session to guide them on how to use the prototype application. One of the information given to the students and teachers was that *VeFRA* allows only multiple choice questions.

## BENEFITS OF VEFRA

The *VeFRA* system offers numerous benefits:

- The *VeFRA* application can reduce the related expenses for the traditional face-to-face learning method which is common in most developing countries such as lecture halls and other learning delivery facilities associated with physical presence between the students and teachers.
- The *VeFRA* application can provide economies of scale at tertiary institutions, as cost of each additional learner is negligible once the lecture materials have been developed and hosted in a central server.
- In solving the digital illiteracy problems common to developing countries, the *VeFRA* application can improve accessibility to education.
- The able (non-physically challenged) is offered an alternative learning platform; and there will be drastic reduction in the illiteracy level since mobile phone which is widely used device all over the world can be used for learning (m-Learning).

## VEFRA EVALUATION

The *VeFRA* application was evaluated for 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 (Ikhu-Omoregbe, 2010). The usability of the E-Learning application was measured to specify the features and attributes required to make the product usable using ISO's standard of usability (ISO 9241-11, 1998).

*Figure 13. A screen shot of tutorial questions web interface*

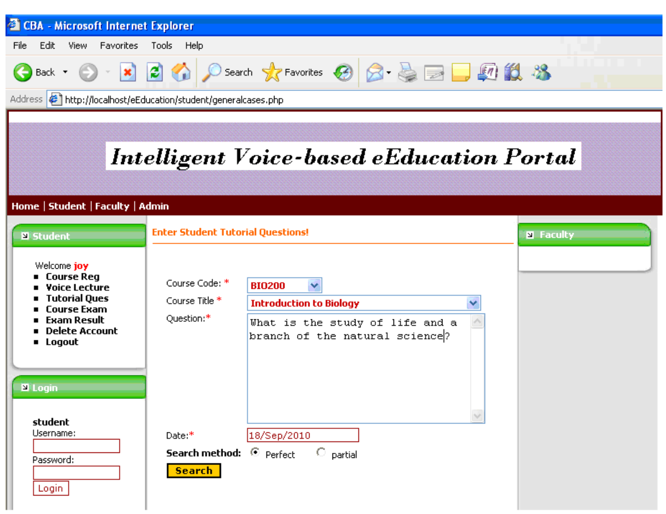
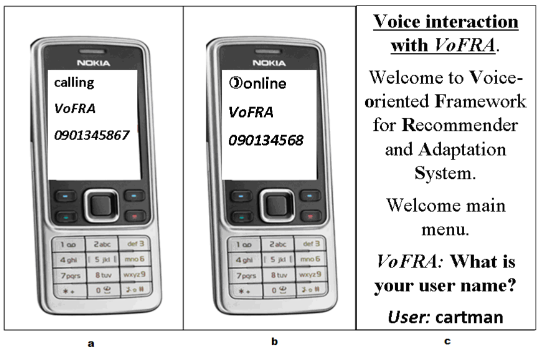


*Figure 14. A model of mobile phone interaction with VeFRA*




## RESEARCH DESIGN

In conducting the usability evaluation, questionnaire were designed and administered. The teachers and students from a school for the blind and partially sighted in Lagos, Nigeria were some of the respondents. The survey questionnaire was divided into five sections namely: background information, user experience with mobile phone and the system, effectiveness of the system, efficiency of the system and user satisfaction with the system.

Each section of Effectiveness, Efficiency and User satisfaction contains five Questions represented by Q1, Q2, Q3, Q4 and Q5. The questionnaire aims at eliciting information from the school in order to measure the usability of the voice-based E-Learning application provided. A total of 70 questionnaires were administered on teachers and students but only 63 responses were received, analyzed and reported (see Table 1). The questions were designed using five-point likert-scale where 1= strongly disagree, 2 = disagree, 3= undecided, 4 = agree and 5= strongly agree.

## EVALUATION RESULTS

An overall score of all the learners was computed for each of the usability dimensions 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 use a mobile phone to dial a number that will connect the learners to the application and show how to navigate within the application. In terms of ownership/usage of mobile phones, users in our sample own/use mobile phones for almost 6 years which resulted to 35 as those having and 28 as those not having experience. In the area of acquiring ICT training to be able to use the application, 51 out of 63 said they would need training. The average effectiveness is 3.98, average efficiency is 4.29 and average users' satisfaction is 4.13.

Some learners were concerned with the time it took for *VeFRA* to recognize a human voice as a result of noise distortion, and that accounted for the low score of 3.98 out of 5 in the question that say "*The voice request & response I got from the*

*Table 1. Descriptive statistical analysis of questionnaire data*

| **Summary of Study Variable** | | | | | | |
|---|---|---|---|---|---|---|
| **Usability Measures** | **Responses from Teachers** | | **Responses from Students** | | **Total # of Respondents** | **Total Mean Rating** |
| | **Visually Impaired Teachers** | **Normal Teachers** | **Visually Impaired Students** | **Normal Students** | | |
| Effectiveness | 5 | 5 | 47 | 6 | 63 | 3.98 |
| Efficiency | 6 | 6 | 47 | 6 | 63 | 4.29 |
| Satisfaction | 6 | 6 | 47 | 6 | 63 | 4.13 |

*Table 2. User survey results (N=63)*

| Survey Questions | | AVG | SD | VAR |
|---|---|---|---|---|
| **Background Information** | | | | |
| Do you own/use a mobile phone? [yes] [no] | | 35/28 | | |
| Would you be able to afford a mobile phone to call the e-Learning application? [yes] [no] | | 39/24 | | |
| Would you support the use of mobile phone for e-Learning in your school? [yes] [no] | | 42/21 | | |
| Do you need more computing skills/training/time to be able to use the system? [yes] [no] | | 51/12 | | |
| **Effectiveness** | | | | |
| Q1 | I was able to complete my task successfully and correctly using the application | 4.00 | 0.92 | 0.84 |
| Q2 | The system did not show error message(s) while using it. | 4.02 | 0.85 | 0.73 |
| Q3 | I was able to recover from my mistakes easily. | 4.16 | 0.83 | 0.68 |
| Q4 | I feel comfortable using the application. | 4.21 | 0.83 | 0.68 |
| Q5 | The voice request & response I got from the system was clear | 3.49 | 0.56 | 0.32 |
| Average Effectiveness | | 3.98 | 0.80 | 0.65 |
| **Efficiency** | | | | |
| Q1 | Using the system saves me time of learning | 4.08 | 0.81 | 0.65 |
| Q2 | I was able to complete my task on time | 4.21 | 0.72 | 0.52 |
| Q3 | I was well able to navigate the voice user interface on time when using the system | 4.41 | 0.74 | 0.55 |
| Q4 | I didn't have to carry out too many/difficult steps before completing my Task | 4.35 | 0.74 | 0.55 |
| Q5 | The learning content I requested for suited my needs | 4.38 | 0.77 | 0.59 |
| Average Efficiency | | 4.29 | 0.75 | 0.57 |
| **User Satisfaction** | | | | |
| Q1 | The system was easy to learn | 4.05 | 0.83 | 0.69 |
| Q2 | The system was easy to use and user friendly | 4.13 | 0.75 | 0.56 |
| Q3 | I am satisfied using the system | 4.11 | 0.84 | 0.71 |
| Q4 | I feel the system met my need | 4.21 | 0.79 | 0.62 |
| Q5 | I am satisfied with the performance of the system in accomplishing my tasks | 4.13 | 0.79 | 0.62 |
| Average User Satisfaction | | 4.13 | 0.80 | 0.64 |
| **Learnability** | | | | |
| Q1 | The system provides clarity of wordings. | 4.10 | 0.82 | 0.67 |
| Q2 | The grouping and ordering of menu options is logical for easy learning. | 4.06 | 0.76 | 0.58 |
| Q3 | The command names are meaningful. | 4.05 | 0.83 | 0.69 |
| Q4 | I could perform tasks on a proficient level as a first time user | 4.13 | 0.81 | 0.66 |
| Q5 | As new user, I was able to orient myself with the system | 4.02 | 0.79 | 0.63 |
| Average Learnability | | 4.07 | 0.80 | 0.65 |

*continued on following page*

*Table 2. Continued*

| Survey Questions | | AVG | SD | VAR |
|---|---|---|---|---|
| **Memorability** | | | | |
| Q1 | I gain the same level of skill as the last time I used the system | 4.02 | 0.81 | 0.66 |
| Q2 | I was able to remember the commands I performed last time | 4.13 | 0.73 | 0.53 |
| Q3 | I was able to reestablish my skills after a long time of using the system | 4.11 | 0.76 | 0.58 |
| Q4 | I could recall every commands I performed moments earlier | 4.14 | 0.76 | 0.58 |
| Q5 | I could remember the ordering of menu structure after a long time of using the system | 4.13 | 0.73 | 0.53 |
| Average Memorability | | 4.11 | 0.76 | 0.58 |
| Resultant Average Rating | | 4.12 | 0.78 | 0.62 |

*system was clear*". The recognition rate of Voxeo ASR engine was ineffective some of the times arising from noise interference. The VoiceXML interpreter for TTS had problems compiling strings stored in the database with single and double quotes ('and"). The quotes were edited out of the text to record success in TTS translation. Notwithstanding, the user survey results were encouraging, particularly, the Question two (Q2) under efficiency in Table 2 that asked '*I was able to complete my task on time*' resulting to 4.21.

## FREQUENCY DISTRIBUTION OF VEFRA USABILITY EVALUATION

In the five questions that make up the efficiency questionnaire category, question five (Q5) contains the question "*The learning content I requested for suited my needs*". This section presents the statistical analysis of Q5 as shown in Figure 15, 16 and 17.

The usability attributes analysis of the evaluation for Effectiveness, Efficiency and User Satisfaction is shown in Figure 18.

## USABILITY EVALUATION

Several usability studies suggest the system with "Very Bad Usability" should have 1 as mean rating, "2 as Bad Usability", 3 as Average Usability, "4 as Good Usability" and "5 as Excellent Usability". Sauro et al., (2005) proposed that "Good Usability" should have a mean rating of 4 on a 1-5 scale and 5.6 on a 1-7 scale. Therefore, it can be concluded that the prototype application developed for the school has "Good Usability" based on the average (AVG) total rating of 4.13.

*Figure 15. Mean rating for Q5 (efficiency)*

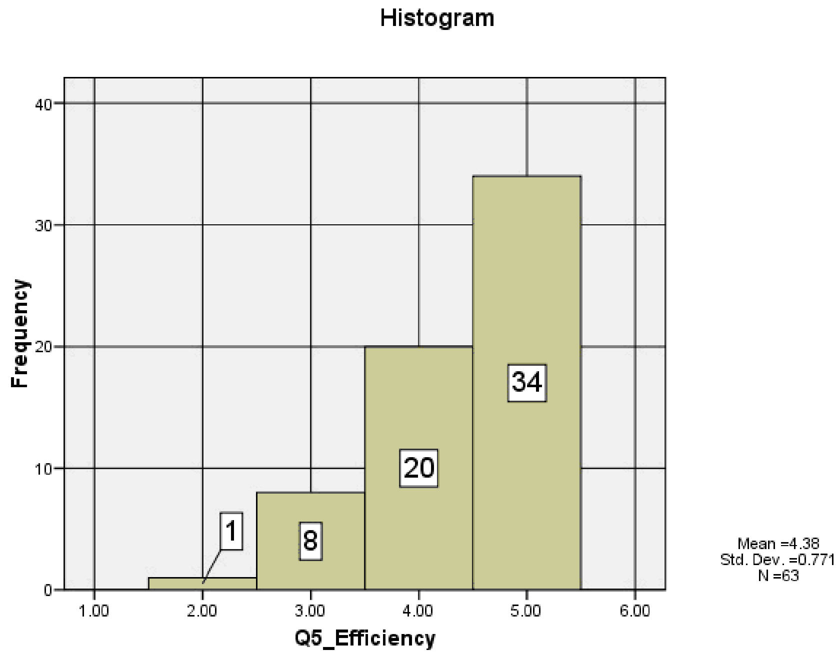
| | | |
|---|---|---|
| N | Valid | 63 |
| | Missing | 0 |
| Mean | | 4.3810 |
| Std. Deviation | | .77102 |
| Variance | | .594 |

*Figure 16. Frequency distribution for Q5 (efficiency)*

Q5_Efficiency

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 2.00 | 1 | 1.6 | 1.6 | 1.6 |
| | 3.00 | 8 | 12.7 | 12.7 | 14.3 |
| | 4.00 | 20 | 31.7 | 31.7 | 46.0 |
| | 5.00 | 34 | 54.0 | 54.0 | 100.0 |
| | Total | 63 | 100.0 | 100.0 | |

*Figure 17. Frequency data for Q5 (efficiency) histogram*

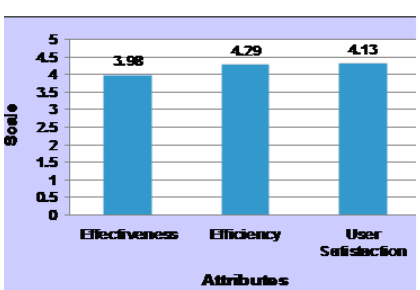



## FUTURE RESEARCH DIRECTIONS

The future research direction, for this study is two-fold. First, evaluation of the system using PARAdigm for Dialogue System Evaluation (PARADISE), and second voice biometric techniques based on speech data will be added as additional means of security mechanism to enhance the authentication of candidates for examination. One key identification technique is the use of voiceprint of the caller. An improved system should therefore be provided as further research to combine the conventional pin/telephone number authentication and biometric to increase security of the system.

With these additional features, the voice learning application will be able to provide a more participatory and secured voice-based educational experience for the students.

*Figure 18. Usability attributes analysis of VeFRA*




## CONCLUSION

In this thesis, a voice-enabled famework for recommender and adaptation system has been provided to offer solution to the problem encountered by learners who use telephone to access Web learning content. This study attempts to solve the problem of using telephone access to find content suitability (FCS) among the numerous learning resources on the Web. The *VeFRA* framework was used as a generic work guideline to develop a prototype intelligent voice-based E-Learning application for the *VeFRA* project. The intelligent component services of the application were realized using CBR, Modified Porter Stemming (MPS) algorithm and ontology. The prototype application was tested in a school for the blind, and the result of evaluation shows that the prototype application developed for the school has "Good Usability" rating of 4.13 out of 5 scale. 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-Learning 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 could equally be useful to students with reading or spelling difficulties (e.g. dyslexia). It is also vital in the area of ubiquitous learning. The E-Learning application will be useful especially for students who are physically challenged such as the visually impaired. It will also be useful for people who have medical history of reacting to Repetitive Strain Injury (RSI) as a result of their sedentary lifestyle using the keyboard. Therefore this system gives every learner equal right and access to quality education irrespective of his or her physical disability.

The E-Learning system offers an alternative platform of learning for learners without disability. It complements the existing E-Learning systems such as Web-based learning, m-Learning, etc. The E-Learning system offers the possibility to learn anytime and anywhere there is telephone access, regardless of the availability of Internet services.

## REFERENCES


Abdel-Badeeh, M. S., & Rania, A. (2005). A case based expert system for supporting diagnosis of heart diseases, *AIML Journal, 5*(1).

Armstrong, H. (2009). Advanced IT education for the vision impaired via e-learning. *Journal of Information Technology Education*, *8*, 244–256.

Aroyo, L., & Dicheva, D. (2004). The new challenges for e-learning: The educational semantic web. *Journal of Educational Technology & Society*, *7*(4), 59–69.

Azeta, A. A., Ayo, C. K., Atayero, A. A., & Ikhu-Omoregbe, N. A. (2008c). Development of a telephone-based e-learning portal. In *Proceedings of the 1st International Conference on Mobile Computing, Wireless Communication, E-Health, M-Health and TeleMedicine (FICMWiComTelHealth'08)*, (pp. 141-149). Ogbomosho, Nigeria: FICMWiComTelHealth.

Azeta, A. A., Ayo, C. K., Atayero, A. A., & Ikhu-Omoregbe, N. A. (2009b). Application of voiceXML in e-learning systems. In *Cases on Successful E-Learning Practices in the Developed and Developing World: Methods for the Global Information Economy*. Hershey, PA: IGI Global. doi:10.4018/978-1-60566-942-7.ch007.

Azeta, A. A., Ayo, C. K., Atayero, A. A., & Ikhu-Omoregbe, N. A. (2009c). A case-based reasoning approach for speech-enabled e-learning system. In *Proceedings of 2nd IEEE International Conference on Adaptive Science & Technology (ICAST).* Accra, Ghana: IEEE. Retrieved from http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5375737

Bobbie, P. O., Fordjour, I., Aboagye, D. O., Dzidonu, C., Darkwa, O., & Nyantakyi, K. (2008). Technology enablers for collaborative virtual education. In *Proceedings of 3rd International Conference on ICT for Development, Education and Training AICC.* Accra, Ghana: AICC.

Bures, I. M., & Jelinek, I. (2004). Description of the adaptive web system for e-learning. In *Proceedings of IADIS International Conference E-Society 2004,* (Vol. 2). IADIS. Retrieved from http://www.iadis.net/dl/final_uploads/200402C042.pdf

Chin, C. C., Hock, G. T., & Veerappan, C. M. (2006). VoiceXML as solution for improving web accessibility and manipulation for e-education. In *Proceedings of School of Computing and IT, INTI College.* Malaysia: INTI. Retrieved from http://intisj.edu.my/INTISJ/InfoFor/StaffResearch/10.pdf

Cranefield, S., Hart, L., Dutra, M., Baclawski, K., Kokar, M., & Smith, J. (2002). UML for ontology development. *The Knowledge Engineering Review*, (17): 61–64.

Devedzic, V. (2004). Education and the semantic web. *International Journal of Artificial Intelligence in Education*, *14*, 9–65.

Dicheva, D., Sosnovsky, S., Gavrilova, T., & Brusilovsky, P. (2005). Ontological web portal for educational ontologies. In *Proceedings of International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL).* Amsterdam: SW-EL. Retrieved from http://www.win.tue.nl/SW-EL/2005/swel05-aied05/proceedings/4-Dicheva-final-full.pdf

Divitini, M., Haugalokken, O., & Morken, E. M. (2005). Blog to support learning in the field: Lessons learned from a fiasco. In *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT '05)*, (pp. 219- 221). IEEE.

Donegan, M. (2000). BECTA voice recognition project report. *In Proceedings of BECTA*. Retrieved from http://www.becta.org.uk/teachers/teachers.cfm?section=2&id=2142

Ermalai, I. And one, D., & Vasiu, R. (2009). Study cases on e-learning technologies used by universities in romania and worldwide. *WSEAS Transactions on Communications, 8*(8), 785–794. Retrieved from http://www.wseas.us/E-library/transactions/communications/2009/29-640.pdf

Farinazzo, V., Salvador, M., & Luiz, A. S. Kawamoto, & de Oliveira Neto, J. S. (2010). An empirical approach for the evaluation of voice user interfaces. *User Interfaces*. Retrieved from www.intechopen.com/download/pdf/pdfs_id/10804

Gallivan, P., Hong, Q., Jordan, L., Li, E., Mathew, G., & Mulyani, Y. … Tappert, C. (2002). VoiceXML absentee system. In *Proceedings of MASPLAS'02: The Mid-Atlantic Student Workshop on Programming Languages and Systems*. Retrieved from http://csis.pace.edu/csis/masplas/p10.pdf

Garcıa, V. M. A., Ruiz, M. P. P., & Perez, J. R. P. (2010). Voice interactive classroom, a service-oriented software architecture for speech-enabled learning. *Journal of Network and Computer Applications*, *33*, 603–610. doi:10.1016/j.jnca.2010.03.005.

Gerber, E., & Kirchner, C. (2001). Who's surfing? Internet access and computer use by visually impaired youths and adults. *Journal of Vision Impairment and Blindness*, *95*, 176–181.

Ghaleb, F. F. M., Daoud, S. S., Hasna, A. M., Jaam, J. M., & El-Sofany, H. F. (2006). A web-based e-learning system using semantic web framework. *Journal of Computer Science*, *2*(8), 619–626. doi:10.3844/jcssp.2006.619.626.

Gil, A., & García-Penalvo, F. J. (2008). Learner course recommendation in e-learning based on swarm intelligence. *Journal of Universal Computer Science*, *14*(16), 2737–2755.

*Gloss*. (2004). Retrieved from http://www.abc.net.au/pipeline/radio/programs/gloss2.htm

Heiyanthuduwage, S. R., & Karunaratne, D. D. (2006). A learner oriented ontology of metadata to improve effectiveness of learning management systems. In *Proceedings of the Third International Conference on E-Learning for Knowledge-Based Society*. Bangkok, Thailand: IEEE.

Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2005). Adaptive e-learning content generation based on semantic web technology. In *Proceedings of Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL).* Amsterdam: SW-EL. Retrieved from http://www.win.tue.nl/SW-EL/2005/swel05-aied05/proceedings/5-Holohan-final-full.pdf

Holzinger, A., Smolle, J., & Reibnegger, G. (2006). Learning objects (LO): An object-oriented approach to manage e-learning content. In *Encyclopedia of Information in Healthcare & Biomedicine*. Hershey, PA: IGI Global.

Ikhu-Omoregbe, N. A. (2010). *Development of a formal framework for usable operations support in e-health-based systems*. (Doctoral dissertation). Covenant University, Ota, Nigeria.

ISO 9241-11. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*. Geneva, Switzerland: ISO.

Jain, L. C., & Martin, N. M. (1998). *Fusion of neural network, fuzzy systems, and genetic algorithms: Industrial applications*. Boca Raton, FL: CRC Press International.

Jianhua, W., Long, Z., Jun, Z., & Xiping, D. (2009). A voiceXML-based mobile learning system and its caching strategy. In *Proceedings of MLearn 2009, 8th World Conference on Mobile and Contextual Learning*. Orlando, FL: IEEE.

Kolias, C., Kolias, V., & Anagnostopoulos, L. (2004). *A pervasive wiki application based on voiceXML.* Retrieved from http://www.icsd.aegean.gr/publication_files/conference/275294916.pdf

Kondratova, I. (2009). Multimodal interaction for mobile learning. In Proceedings *of the 5th International Conference (UAHCI '09).* San Diego, CA: UAHCI.

KSL. (2005). *KSL software and network services*. Retrieved from http://www.ksl.stanford.edu/sns.shtml

Larson, J. (2000). *Introduction and overview of W3C speech interface framework*. W3C Working Draft. Retrieved from http://www.w3.org/TR/2000/WD-voice-intro-20001204/

Lerlerdthaiyanupap, T. (2008). *Speech-based dictionary application*. (MSc Thesis). University of Tampere, Tampere, Finland.

Lin, W. J., Yueh, H. P., Liu, Y. L., Murakami, M., Kakusho, K., & Minoh, M. (2006). Blog as a tool to develop e-learning experience in an international distance course. In *Proceedings of the Sixth International Conference on Advanced Learning Technologies (ICALT'06).* ICALT.

Motiwalla, L. F. (2009). A voice-enabled interactive services (VòIS) architecture for e-learning. *International Journal on Advances in Life Sciences, 1*(4), 122–133. Retreived from http://www.iariajournals.org/life_sciences/lifsci_v1_n4_2009_paged.pdf

Motiwalla, L. F., & Qin, J. (2007). Enhancing mobile learning using speech recognition technologies: A case study. In *Proceedings of the Eighth World Congress on the Management of eBusiness (WCMeB'07).* WCMeB.

Niels, O. B. (2000). *Draft position paper for discussion at the ELSNET brainstorming workshop.* Retrieved from http://www.elsnet.org/dox/rM-bernsen-v2.pdf

Nisbet, P. D., & Wilson, A. (2002). *Introducing speech recognition in schools: Using dragon natural speaking*. Edinburgh, UK: University of Edinburgh.

NUC. (2004). The state of Nigerian universities. *Nigerian University System Newsletter, 2*(1).

Oboko, R., Wagacha, P. W., & Omwenga, E. (2008). Adaptive delivery of an object oriented course in a web-based learning environment. In *Proceedings of 3rd International Conference on ICT for Development, Education, and Training*. Accra, Ghana: IEEE.

Ondas, I. S. (2006). VoiceXML-based spoken language interactive system. In *Proceedings 6th PhD Student Conference and Scientific and Technical Competition of Students of Faculty of Electrical Engineering and Informatics Technical University of Košice.* Košice, Slovakia: IRKR. Retrieved from http://irkr.tuke.sk/publikacie/_vti_cnf/Prispevok_eng.pdf

Pakucs, B. (2003). 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: EACL.

Paul, D. (2003). *Speech recognition for students with disabilities*. Edinburgh, UK: University of Edinburgh.

Peter, Y., Vantroys, T., & Lepretre, E. (2008). Enabling mobile collaborative learning through multichannel interactions. In *Proceedings of 4th International Conference on Interactive Mobile and Computer Aided Learning (IMC'08).* IMC.

Raghuraman, M. B. (2004). *Design and implementation of V-HELP system–A voice-enabled web application for the visually impaired.* (Unpublished Master Thesis). University of Nebraska, Lincoln, NE.

Sauro, J., & Kindlund, E. (2005). A method to standardize usability metrics into a single score. In *Proceedings of CHI'05*. Portland, OR: ACM.

Spasic, I., Ananiadou, S., & Tsujii, J. (2005). *MaSTerClass: A case-based reasoning system for the classification of biomedical terms*. Retrieved http://qr.net/jzsU, accessed 2005.07.05

*TellMe Voice Portal*. (2002). Retrieved from http://www.tellme.com

Voice, X. M. L. (2007). *VoiceXML application development life cycle.* Palo Alto, CA: Hewlett Packard Development Company. Retrieved from staff.washington.edu/benroy/ivr/vx_devlifecycle.pdf

Voxeo. (2003). *Voice voice server*. Retrieved from http://community.voxeo.com

Walsh, P., & Meade, J. (2003). Speech enabled e-learning for adult literacy tutoring. In *Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03).* IEEE. Retrieved from http://i-learn.uitm.edu.my/resources/journal/j1.pdf

Williamson, K., Wright, S., Schauder, D., & Bow, A. (2001). The internet for the blind and visually impaired. *Journal of Computer Mediated Communication*. Retrieved from http://jcmc.indiana.edu/vol7/issue1/williamson.html

Zaiane, O. R. (2005). Recommended systems for e-learning: Towards non-intrusive web mining. *Data Mining in E-Learning, 2.*

Zaykovskiy, D. (2006). Survey of the speech recognition techniques for mobile devices. In *Proceedings of Speech and Computer (SPECOM'06).* St. Petersburg, Russia: SPECOM.

## ADDITIONAL READING

Aarnio, T. (1999). *Speech recognition with hidden markov models in visual communication*. (Unpublished Master's Thesis). University of Turku, Turku, Finland.

Abbott, K. (2002). *Voice enabling web applications: VoiceXML and beyond*. Berkeley, CA: Apress. doi:10.1007/978-1-4302-0850-1.

Abraham, A. (2005). *Handbook of measuring system design*. New York: John Wiley & Sons.

Abu-Zaina, H., Kirk, D., & Elmer, A. (1996). A speech recognition system for operation of a visual inspection workstation. In *Proceedings of RESNA'96*. Wichita, KS: RESNA. Retrieved from http://www.dinf.ne.jp/doc/english/Us_Eu/conf/resna96/page143.htm.

Adaptive. (2005). *Adaptive and assistive technologies in e-learning*. Retrieved from http://www.adcet.edu.au/StoredFile.aspx?id=1353&fn=adptve_assstv_techn_elearning.Pdf

Begum, S., Ahmed, M. U., Funk, P., Xiong, N., & Schéele, B. B. (2007). Similarity of medical cases in health care using cosine similarity and ontology. In *Proceedings of The 5th Workshop on CBR in the Health Sciences*. Belfast, Ireland: ICCBR. Retrieved from www.cbr-biomed.org/workshops/ICCBR07/iccbr07wk-Begum.pdf

Belkin, N. J. (2005). *Intelligent information retrieval: Whose intelligence. Understanding and Supporting Multiple Information Seeking Strategies*. Newark, NJ: Rutgers University.

Burke, D. (2007). *Speech processing for IP networks, media resource control protocol (MRCP)*. New York: John Wiley & Sons. doi:10.1002/9780470060599.

Chang, S. E. (2007). An adaptive and voice-enabled pervasive web system. *International Journal of Enterprise Information Systems*, *3*(4), 69–83. doi:10.4018/jeis.2007100105.

Chowdhary, K. R. (2005). *Current trends in information retrieval.* Retrieved from http://ncsi-net.ncsi.iisc.ernet.in/gsdl/collect/icco/index/assoc/HASH014f.dir/doc.doc

Cisco. (2010). *Cisco unified customer voice portal 8.0.* Retrieved from http://qr.net/imicsn4a

Cook, S. (2002). *Speech recognition HOWTO.* Retrieved from http://tldp.org/HOWTO/Speech-Recognition-HOWTO/software.html

Divakaran, D. (2011). *Create and setup a website in 10 mins.* Retrieved from http://www.deepeshmd.com/wp-content/uploads/CreatE-and-Setup-a-WebsitE-in-10-mins.pdf

Ehsani, F., & Knodt, E. (1998). Speech technology in computer-aided language learning: Strengths and limitations of a new call paradigm. *Language Learning & Technology*, *2*(1), 54–73. Retrieved from http://llt.msu.edu/vol2num1/article3/.

Fabbrizio, G. D., Okken, T., & Wilpon, J. G. (2009). A speech mashup framework for multimodal mobile services. In *Proceedings of ICMI-MLMI'09*. Cambridge, MA: ICMI. Retrieved from http://www.difabbrizio.com/papers/icmi174-difabbrizio.pdf

Farinazzo, V. Salvador1, M., Luiz, A. S., Kawamoto, & de Oliveira Neto, J. S. (2010). An empirical approach for the evaluation of voice user interfaces. *User Interfaces*. Croatia: INTECH. Retrieved from www.intechopen.com/download/pdf/pdfs_id/10804

Fiedler, G., & Schmidt, P. (2004). *Developing interactive voice response interfaces for large information systems*. Retrieved from http://www.mgovernment.org/resurces/euromgov2005/PDF/19_S042FG-S04.pdf

Freitas, D., & Kouroupetroglou, G. (2004). Speech technologies for blind and low vision persons. *Technology and Disability*, *20*, 135–156.

Hakulinen, J. (2006). *Software tutoring in speech user interfaces*. (Doctoral Dissertation). Dissertations Inc.

Hisayoshi, I., & Toshio, O. (1999). *Case based reasoning system for information technology education*. Institute of Electronics, Information and Communication Engineers.

HKU. (n.d.). *Development of the HKU campus portal*. University of Hong Kong. Retrieved from http://www.hku.hk/cc_news/ccnews100/portal.htm

Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2005). Adaptive e-learning content generation based on semantic web technology. In *Proceedings of International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL)*. Amsterdam: SW-EL. Retrieved from http://www.win.tue.nl/SW-EL/2005/swel05-aied05/proceedings/5-Holohan-final-full.pdf

Howard, S. (2006). *Interactive training on mobile devices: Training using text-to-speech and speech recognition, on pocket PC PDA*. (Unpublished Master's Thesis).

Husni, H., & Jamaludin, Z. (2009). ASR technology for children with dyslexia: Enabling immediate intervention to support reading in Bahasa Melayu. *US-China Education Review*, *6*(6), 64–70.

Jankowska, B. (2007). *Architectural frameworks for automated content adaptation to mobile devices based on open-dource technologies*. (Unpublished Doctoral Thesis). European University Viadrina, Frankfurt, Germany. Retrieved from http://opus.kobv.de/euv/volltexte/2007/23/pdf/PHD_BJankowska.pdf

Jumar, O. Cizmar, & Rusko, M. (2006). Galaxy/VoiceXML based spoken slovak dialogue system to access the internet. In *Proceedings of ICSLP'06, Ninth International Conference on Spoken Language Processing*. Pittsburgh, PA: ICSLP. Retrieved from http://www.ling.uni-potsdam.de/ws-ecai06/pdf/juhar.pdf

Selouani, S., Le, T., Moghrabi, C., Lanteigne, B., & Roy, J. (2006). Online collaborative learning system using speech technology. In *Proceedings of World Academy of Science, Engineering and Technology Volume*. Retrieved from http://www.waset.org/pwaset/v15/v15-11.pdf

The National Center on Accessible Information Technology in Education. (2006). *AccessIT*. Author.

## KEY TERMS AND DEFINITIONS

**Adaptation:** To adjust the content of objects to suit users' request. Adaptive system monitors particular user's behaviour and characteristics.

**Case-Based Reasoning (CBR):** CBR is a technique for solving new problems based on experiences of previous solutions.

**Electronic-Learning (E-Learning):** E-Learning is the use of Information and Communication Technology (ICT) to support learning processes.

**Framework:** A guideline for achieving an objective.

**Intelligent:** A system is said to be intelligent if it mimics the behavior of human.

**Mobile-Learning (M-Learning):** M-Learning is also known as mobile learning. It is a type of learning that enables the learner to moves from one place to another and learn using a mobile device.

**Recommendation:** Recommendation is the process whereby software agents recommend appropriate options for users.

**Speech:** Speech is used interchangeably with voice. A speech is a sound signal used for language communication. Superficially, the speech signal is similar to a sound produced by a musical instrument, although it is more flexible and varied. When we speak, we push air from our lungs through the vocal chords, sometimes tightening the chords to make them vibrate as the air passes over them.

**VoiceXML:** VoiceXML is an extension of the Extensible Markup Language (XML) that allows interactive access to the Web through the telephone or a voice browser.