



## International Journal of Advance Research Publication and Reviews

Vol 2, Issue 3, pp 40-47, March 2025

# Development of a Web-Based Language Learning System Using an Adaptive Voice Recognition Paradigm(A Case Study of Etuno Language Basic Words)

*Olowookere, Abiodun S.<sup>1</sup>; Ayangbekun, Oluwafemi J.<sup>2\*</sup>; Adeola, Olajide O.<sup>3</sup>; Adams, Racheal A.<sup>4</sup>*

<sup>1,3,4</sup>Department of Computer Science, Oyo State College of Agriculture and Technology, Igboora, Oyo State, Nigeria.

<sup>2</sup>Department of Computer Engineering, Federal Polytechnic Ayede, Ogbomoso, Oyo State, Nigeria.

\*ayangbekunoj@gmail.com

### ABSTRACT

This language learning system with speech recognition is designed to help individuals improve on already learnt language(s) or learn a new one in an interactive and engaging way. The system also features real-time feedback through transcription, allowing learners to adjust their pronunciation and intonation of Etuno language as they practice. With its user-friendly interface and comprehensive learning resources, this language learning system with speech recognition is an effective tool for individuals looking to master the desired language. Learners are able to track their progress and receive feedback based on the transcription for further study, based on their performance. An adaptive voice recognition language learning system as an innovative tool designed to help language learners improve their speaking skills. One of the main benefits of an adaptive voice recognition language learning system is that it can provide immediate feedback to the learner through the transcription. This allows learners to quickly identify their mistakes and correct them, leading to faster progress and improved speaking skills. Overall, an adaptive voice recognition language learning system is an effective and efficient tool for language learners who want to improve their speaking skills. Learners can practice speaking with confidence and accuracy.

**Keywords:** Adaptive, Etuno, Language, Learning, Recognition.

### 1. Introduction

Learning is one of the very vital human activities that require concentration coupled with interactivity, clear and distinct understanding of the facts being stated or discussed. It also involves high communication skills and techniques, attractive learning qualities such as colorful pictorial presentation of information (otherwise known as Visual Representation) among others. Nevertheless, not all learning process is considered to be effective. Effectiveness can be assumed to have been achieved if about hundred percent (100%) of the population lectured assimilate the knowledge and/or ideas being passed across; and this is quite different when the population size is very high and there is existence of certain un conducive learning conditions such as noise, poor ventilation, and extreme temperatures.

Computer-based learning also known as computer-aided instruction makes use of the interactive elements of the computer applications and software and the ability to present any type of media to the users. Computer-based learning has many benefits, including the advantage of users learning at their own pace and also learning without the need for an instructor to be physically present. It provides more learning opportunity for people from disadvantaged environments. In respect to scope of this paper, users will learn easily with the help of artificial intelligence (Speech Recognition) and allow users to get impacted through technology with less cost and stress is achieved. The learning also offers safety and flexibility as well as helps learners to track their progress by reduction of overall training time. As speech recognition is the dominant means for communication between humans, it promises to be important for communication between humans and machines, if it can just be made a little more reliable. In Speech recognition, the acoustic signal is being converted to a set of words. The applications include voice commands and control, data entry, voice user interface, automating the telephone operator's job in telephony, etc. But despite the availability of various language learning resources, many people struggle to effectively learn a new language due to the lack of personalized learning experiences. The current language learning systems are often too rigid, not adapting the individual needs, pace and preferences of the learners. This leads to low motivation, lack of progress and high dropout rates. With the adoption of the speech recognition, learning is made easier for learners such that they do not have to type the word(s) they want translated.

## **2. Literature Review**

### **2.1 Etuno Language**

The name Etuno was derived due to the presence of abundant palm trees "Uno" and "Èté" being the name of the language spoken by the people of the land (i.e., Etuno language). This makes Etuno language being spoken by the Etuno people of Igarra land of Edo State. Meanwhile, Igarra town is the headquarters of Akoko-Edo Local Government Area of Edo State. The Etuno language is 'cousin' to the Yoruba language due to the geographical proximity between Igarra land and the Yorubas. Igarra has no serious linguistic affinity with any other ethnic group in Edo State. The areas of linguistic affinity are the EbiraTotos of Nassarawa state, the Toto town of Plateau State, the Ebira Koto of KotonKarfe of Kogi State as well as the Ebira Tao in Okene also in Kogi State. Other areas with which the Etuno language shares linguistic affinity includes Abuja, Itako, and Ukyia areas of the Federal Capital Territory (Egbokhare, F. O., 2022). .

### **2.2 Learning and Language**

Learning is a process that leads to change, which occurs as a result of experience and increases the potential for improved performance and future learning (Ambrose, S. A. et al, 2019). Learning is the relatively permanent change in a person's knowledge or behavior (Lachman, S. J., 1997) due to experience. Domjan defines learning as an enduring change in the mechanisms of behavior (Domjan, M., 2010). It is a complex process of discovery, collaboration, and inquiry facilitated by language. Language on the other hand is a social and uniquely human means of representing, exploring, and communicating meaning; as well as being a defining feature of culture. It is a means of communication; and it's the primary basis of all communication and the primary instrument of thought. Meanwhile, the relationship between language and culture cannot therefore be overemphasized; as language is a medium through which the culture of a particular society is communicated.

### **2.3 Speech / Voice Recognition**

Research in AI has focused chiefly on the following components of intelligence: learning, reasoning, problem solving, perception, and using language. Speech recognition, also known as automatic speech recognition (ASR), computer speech recognition, or speech-to-text, have capability that enables a program to process human speech into written format. It is capable of converting spoken language (an audio signal) into written text that is often used as a command (Deng L. et. al., 2014).

In 1996, IBM introduced the first voice recognition product that could recognize continuous speech. Voice recognition systems enable users to interact with technology. It is a computer software program or hardware device with the ability to

decode human voice. Voice recognition has gained prominence and use with the rise of AI and intelligent assistants. Automatic speech recognition is one example of voice recognition. Voice recognition software on computers requires that analog audio be converted into digital signals, known as analog-to-digital conversion. For a computer to decipher a signal, it must have a digital database, or vocabulary, of words or syllables, as well as a speedy means for comparing this data to signals. The speech patterns are stored on the hard drive and loaded into memory when the program is run. However, the size of a voice recognition program's effective vocabulary is directly related to the random-access memory capacity of the computer in which it is installed. While voice recognition technology originated on PCs, it has gained acceptance in both business and consumer spaces on mobile devices and in home assistant products. Voice recognition, combined with the growing stable of internet of things sensors, has added a technological layer to many consumer products that previously lacked any smart capabilities. However, voice recognition has grown sporadically in AI and machine learning domain; as several voice recognition software have been implemented to interact with users.

### **3. Methodology**

---

This research is of significant important as it adopted combination of multimedia features, artificial intelligence features, and human-computer interaction principles for the online concept development. The language site will be designed for the learning of the language at its introductory level; while the proposed learning avenue will be a deviation from the traditional teaching and learning procedure. This is to established an effective learning of the subject topic / language; focusing on developing a captivating learning application through the provision of different learning platforms and skillful use of multimedia leading to the enhancement of effective learning and reduction of teachers / tutors teaching workload. Also, the cost of learning materials and duration will be reduced significantly when compared to the cost of traditional / conventional instructional system.

#### **3.1 Adopted Scope**

This paper develops a language learning system that leverages an adaptive learning technique. The system makes use of AI algorithms to analyze the learners performance and adapt the curriculum to their specific needs, abilities and learning pace. The system will also provide personalized feedback reinforcement and motivation to enhance the learning experience and improve outcomes. Meanwhile the ultimate objectives is to make language learning more effective, engaging and accessible for learners at all levels using an interactive presentations. In the course of the system development, security consciousness likes impersonation is being considered while the security measure like password is implemented; with an irresistible user interface to boost its awareness and marketing. Also, the physiological and cognitive impact of the proposed system on their users are addressed using help menu, flash animations, hyperlinks and still images.

#### **3.2 Model Description**

Unlike the traditional learning approach, this new system development introduces an adaptive paradigm for the process of learning language with the help of artificial intelligence to improve the existing system. In the proposed system, users use his/her voice to communicate to the system for it to give results related to voice command he/she gives, and this will improve the efficiency of the system and produce validating results using an Acoustic Model.

##### **3.2.1 Acoustic Model :**

Acoustic modeling of speech typically refers to the process of establishing statistical representations for the feature vector sequences computed from the speech waveform. It also encompasses “pronunciation modeling”, which describes how a sequence or multi-sequences of fundamental speech units (such as phones or phonetic feature) are used to represent larger speech units such as words or phrases which are the object of speech recognition. Acoustic modeling may also include the use of feedback information from the recognizer to reshape the feature vectors of speech in achieving noise robustness in speech recognition (Do Van H., 2015).

Hidden Markov Model (HMM) is one of the most common types of acoustic models. Other acoustic models include segmental models, super-segmental models (including hidden dynamic models), neural networks, maximum entropy models, and (hidden) conditional random fields, etc. Speech recognition engines usually require two basic components in order to recognize speech. One component is an acoustic model, created by taking audio recordings of speech and their transcriptions and then compiling them into statistical representations of the sounds for words. The other component is called a language model, which gives the probabilities of sequences of words. Language models are often used for dictation applications.

### 3.3. Model Flow & Case Diagram

Development of the system was based on the model below (Figure 1). It involves requirement analysis, system design, implementation, testing and validation. This approach below describes the sequence of the steps involved and the associated case diagram (Figure 2).

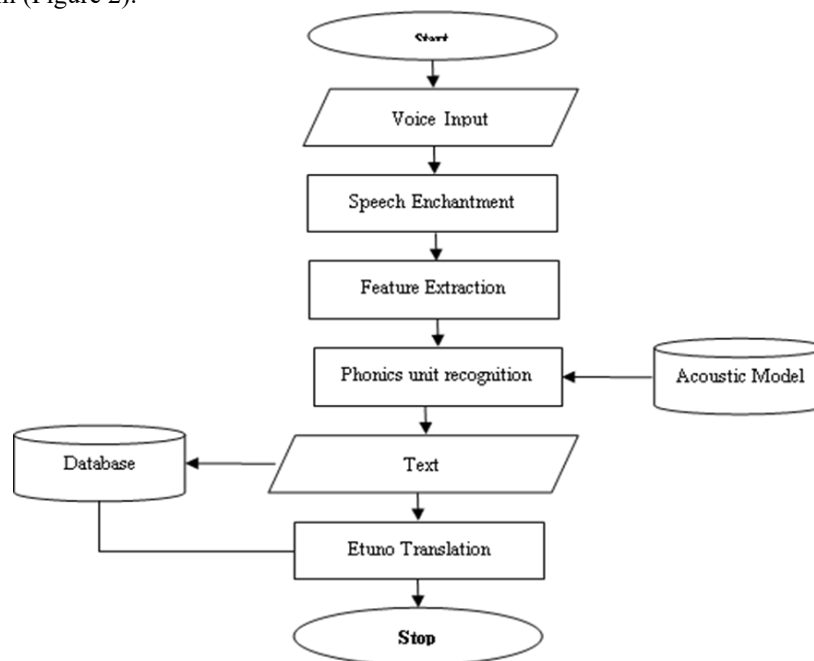


Fig. 1 - Model Flow Diagram

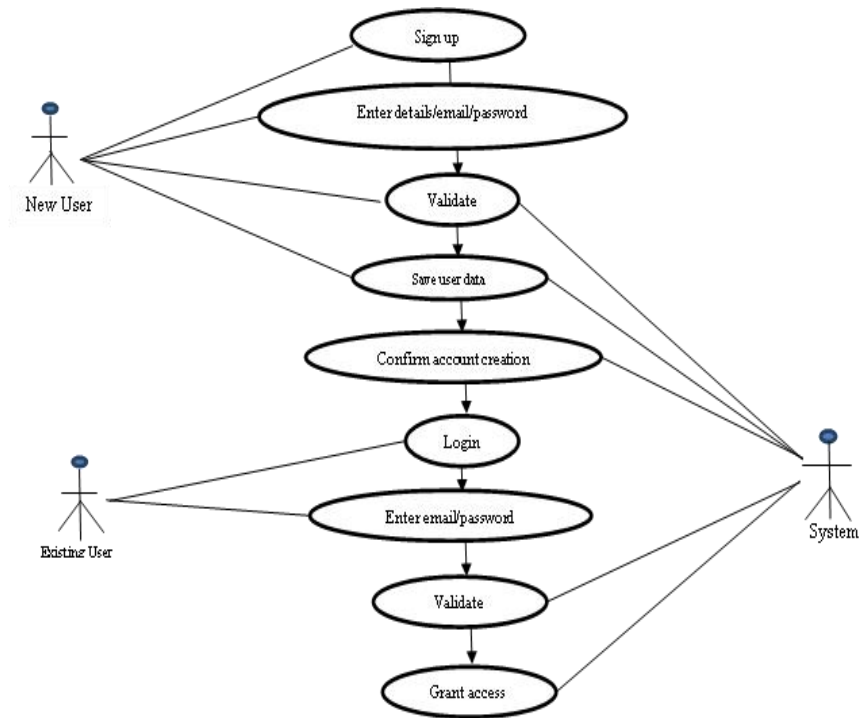


Fig. 2 - Use Case Diagram

#### 4. System Development, Results and Discussion

The system development involves the use of the following programming languages: JavaScript, HTML, CSS and Mongo DB forms. Implementation of the new system requires a high configuration system hardware with the aforementioned Softwares. However, the developed system (Fig. 3) is made up of the following modules: Sign Up; Login; Learning and Speech Recognition.

##### 4.1 System Development

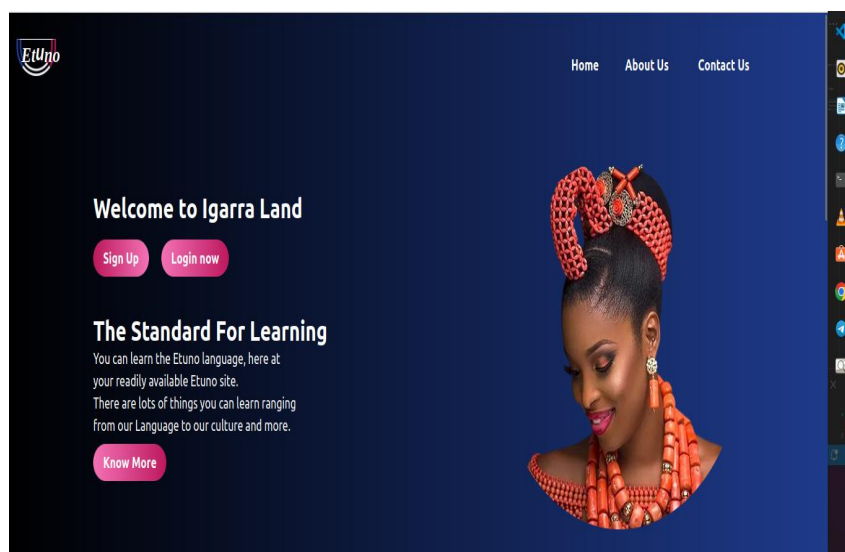


Fig. 3 - System Homepage

## 4.2 Account Creation / Login

The system Home/Welcome Page (Fig. 3) is the first interface that comes up when the system is accessed displaying other interface / links that users interact with in the language context (Etuno Language of Igarra people); such as About Us (Fig. 4) - telling us more about the Igarra people who speak the Etuno language; among others.

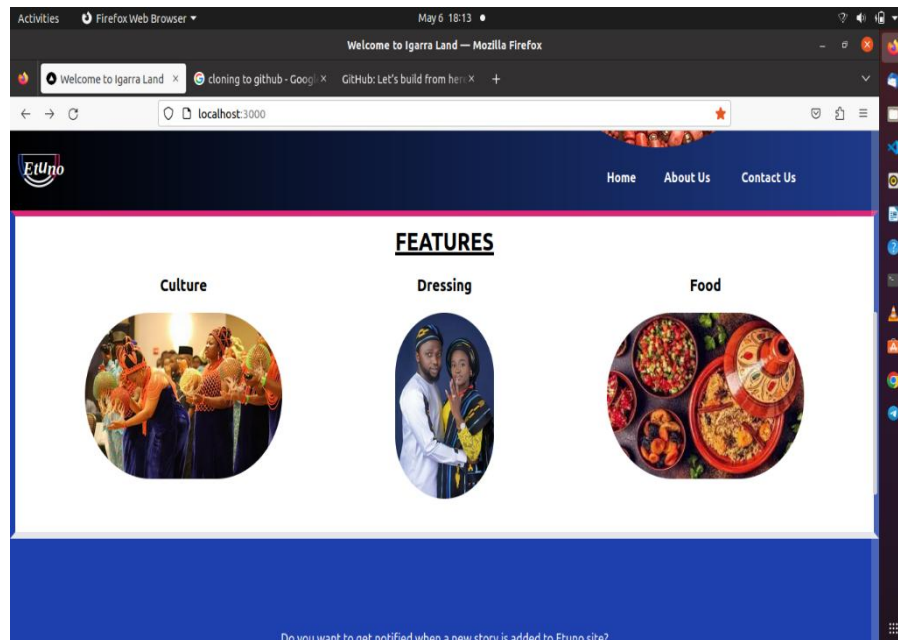


Fig. 4 - About Us Page

## 4.3 Learning module

This interface connects users to Etuno Language learning tutorials involving vowels and consonant sounds / speech and sound to text conversion for easier and seamless learning process (Fig. 5).

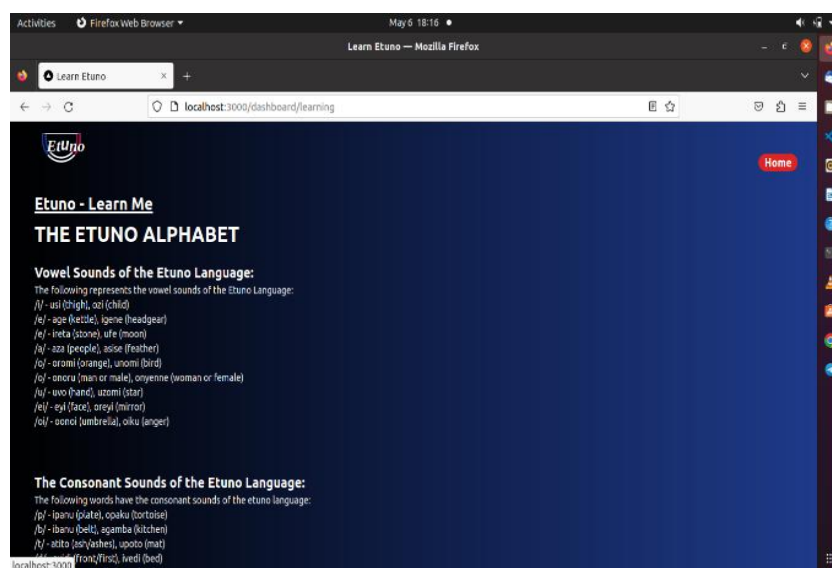


Fig. 5 - Learning Page

#### 4.4 Speech Recognition Module



Fig. 6 - Microphone On / Off Speech to Text

When users click on the start button, the microphone automatically turns on, allowing the system to receive voice inputs from users. However, immediately after users stop speaking, the microphone automatically turns off in order to turn the voice input into text (Fig. 6). Also, when the stop button is clicked, the system can no longer receive voice input.

#### 5. Conclusion

An adaptive voice recognition language learning system is an innovative tool designed to help language learners improve their speaking skills. This system uses advanced speech recognition technology that can adapt to the learner's speech patterns and provide personalized feedback (allows learners to quickly identify their mistakes and correct them, leading to faster progress and improved speaking skills) and correction. The system works by analyzing the learner's speech and comparing it to a native speaker's pronunciation. It can then identify areas where the learner needs improvement and provide specific feedback on how to improve. The system can also adjust the difficulty level of the exercises to match the learner's skill level, ensuring that the learner is always challenged but not overwhelmed. Speech recognition tools greatly enhance learner's skills in a language both when reading and writing. With this technology, learners can obtain information on important elements of phonemic awareness, such as the correspondence between sounds and symbols. As students/learners speak, they see their words on the screen. In conclusion, an adaptive voice recognition language learning system is more becoming an increasingly popular tool in language education as the technology continues to improve.

#### References

- Ambrose, S. A., Norman, M. K., DiPietro, M., Lovett, M. C., & Bridges, M. W. (2019). How learning works: seven research-based principles for smart teaching. Tantor Media, Pp. 301.
- Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and Trends in Signal Processing*, 7(3-4), 197-387.
- Domjan, M. (2010). *Principles of learning and behavior* (6th Ed.). Belmont, CA: Wadsworth/Cengage.
- Do Van H. (2015). *Acoustic Modeling for Speech Recognition under Limited Training Data Conditions*. Nanyang Technological University.

- 
- Egbokhare, F. O. (2022). Unification and Harmonization of Afenmai Languages: Issues and Challenges. *Current Issues in Descriptive Linguistics and Digital Humanities: A Festschrift in Honor of Professor Eno-Abasi Essien Urua*, 333-342.
- Lachman, S. J. (1997). Learning is a process: Toward an improved definition of learning. *The Journal of psychology*, 131(5), 477-480.