# Semantic Speech Annotation in the Ubiquitous Learning Environment

R.Shanmugapriya[1], S.RajaMohammed[2]

PG Scholar, Department of Computer Science and Engineering,  Kalaignar Karunanidhi Institute of Technology,

Coimbatore,  TN, India[1].

Associate professor, Department of Computer Science and Engineering, Kalaignar Karunanidhi Institute of Technology, Coimbatore,  TN, India[2].

**ABSTRACT:** Skill based learning is adopted where the performance of the system is more important including accuracy. So job searchers in medical field based learning, computerized numerically controlled system look for people with high level skills. So that the overall performance of the job could be improved by means of reducing the error distribution across various level by training rigorously with appropriate skill levels. This helps in correcting mistakes and also to access the overall learning process with the better understanding to process this the voice recognition technique is to be used.Speech recognition, which is also referred to as speech to text or voice recognition, in which we recognizing speech as in the Natural Language Processing, and also allowing voice to serve as the "main interface between the human and the computer. This paper mainly focuses on the tools that are used to deliver how current speech recognition technology facilitates the learning part of the students, and also in what way the technology will be help to developing the advance learning system for future. The tools like speech to text conversion are to be discussed in this paper. Although speech recognition has a potential benefit for students with physical disabilities and harsh learning disabilities, these are the technology in which it has been implemented inconsistently in the classroom above the years. By means of this the knowledge continues to develop, on the other hand, various issues are being addressed.

*Keywords:* Speech Recognition, Speech to text, learning Disabilities, Natural Language Processing, knowledge

## I. INTRODUCTION

In this paper we have analyzed that the semantic annotation is one of the part of the skill-based learning environment which provide how the student learn [1]. The main functionality of this is to convert the spoken word into written text. All the speech to text system has two models they are acoustic model and a language model. If we build a large set of model then need to create the pronunciation model. All of these models need only language, dialect, application domain, type of speech, and communication channel. The accuracy of the model is depending upon the user and the environmental condition. The speech which is well formulated only transcribed without any ambiguity. The speech to text conversion system can be categorized based on the use in which the control, dialog system, text dictation, audio document transcription, and so on.

Before the speech tools we need to see about the concept of semantic learning, that is the concept of machine learning in which the system need to understand what the user denotes and also the functionality will be based upon the user speech and the main ultimate aim of the concept is that the system need to recognize by means of learning. Semantic Web is to identify a set of technology, tool and also in standard which form the basic building blocks of a system that could support the vision of a Web imbued with meaning.

The Semantic Web has been developing a layered architecture. The text pre-processing phase is begins with the transformation process of the original unstructured documents. This transformation aims to obtain the desired representation of documents in XML format. The filtered and stemmed XML documents are then index by using the weighting scheme. If the textual data is indexed, either manually or automatically, the indexing structures can be used as a basis for the actual knowledge discovery process. The ubiquitous computing lead to the simple technology in

which it combines more functionality for to use wherever and whenever the learning is possible regardless of various digital devices. It will have limited interaction and technology. Also it will provide learning resources to someone, anytime without any restrictions.

## II. SKILLS-BASED LEARNING ENVIRONMENTS

Active participation by the students in a session will include the live observation of the activity by mentors performing roles within the scenario as well as lecturers observing the session from the control room.

Peer observation of the students in which one collection of students is occupied in a scenario, a second group is observing. Debriefing sessions will be held immediately after the activity debrief about the session will be held, which is facilitated by the lecturers who participated in the scenario.

Then the self-reflection after the event for the students will reflect on the activity and their own performance at a later time. Educator reflection on activities will be of the mentors and control room observers may wish to reflect on the activity in order to assess individual student's performance or refine the scenarios.

The educator reflection across activities will focus on the learners who may wish to consider a cycle of scenarios in terms of their efficiency as a learning tool, or, to address more specific research questions, for example examining how hygiene and infection control approaches are being used by the student cohorts as a whole.

Initially the user will speak out the statement and the system need to convert the spoken statement into text and then need to identify the semantic meaning of the given statement. The annotation here means the metadata functionality of the given sentence or statement the function needs to annotate the system in effective way.

## III.SPEECH RECOGNITION VS SPEECH-TO-TEXT

By researching about the speech recognition tools we may see technologies which referred to as speech-to-text, voice recognition, and speech recognition but in occasionally all within the same product explanation. However the terms can be puzzling, they are all refer to technologies which can translate the spoken language into digitized text. Voice recognition is referring to that the products need to be trained in the way to recognize a specific voice.

The main and ultimate benefits for students by means of disabilities can be also taking in for enhancing the right to make use of the computer, also it will be increased in case of the  writing production, improvements will be in writing procedure, increased freedom, decreased nervousness around writing, and also improvements in core reading and writing abilities.

## IV. TOOLS

In this paper we have explained about various tools which are trained and analyzed based on the accuracy,

*A.  Voxsigma*

The Vox Sigma is the software which is suited for the large set of dictionary which function as speech to text with some extends accuracy. The feature of this tool is, which suits for the professional users in which need to transcribe large set of audio and the video documents for the real time and also for to analyses the call-center data. There are three steps in this software they are given below,

    i.     Identify the audio segments
    ii.    Recognize the spoken language.
    iii.   Finally convert the recognized speech to text.

It has the features in which the noisy speech will be also allowed for to recognize such that no need to have a specific place to give the input. Finally the result will be in xml document which is fully annotated file and also it have speech and non speech segments, labels speaker, words with time code, high quality score with confidence and punctuations. This output file is directly indexed by the search engine. This system is proportional to the system error rate.

### B.  DRAGON SPEECH RECOGNITION SOFTWARE

Dragon is one of the best software in the speech recognition it will change the talk into text and make the process as easy and much faster. The ultimate feature of this tool is to create the manuscript, sending email, and searching in the web, by means of simple voice commands. The accuracy of this tool is up to 99% and also the time taken for this is very less and can get the result in some fraction of time. This can make the computer to control and can check the transformation of the words exactly what we have spoke on the screen. The feature of this tool is also to create documents, write papers, send email, and search the Web by our voice not to spend time to type everything. It is the perfect solution for students, family and others those looking for accurate dictation and to control their home computer by voice. This home based tool (dragon 12) is used to improve the accuracy and also deliver up to 20% when compare to the previous version. A faster processor yields faster performance for 4GB of RAM capacity.

### C.  SPEAKONIA

Speakonia is in reality a diffident freeware service which reads the text out loud. It has around 20 voices in which also have all strong robotic accents and can also lets you rapidly change the reading speed and pitch. The main problem in this tool is that it has to correct the pronunciation of words by means of proper names, for instance. It save the text in WAV files but unfortunately it can't able to save the sound files in the MP3 format, so need to have the third-party alteration utility to convert. But the interface for the program is clear and also straightforward. But this tool is used to function with the text file not to have with the speech recognition.

### D.  TAZTI

Tazti is one of the speech recognition software developed by Voice Tech Group mainly developed for the window personal computers which can support by both 32 and 64 bit versions. The features of this software are that dictation, voice search, PC video game play by voice like command and control input and user configurable speech commands and also in robotics, Google voice search. Also it can able to train a custom voice profile in which allows the user to understand better. It utilize minimal user interface in which the program translate the dictated word after pressing the pause button. This is one disadvantage when compare to the other tools because it won't get the word when the time of speaking so time taken will be high when compare to other tools.

### E.  E-SPEAKING

This tool is also the speech recognition in which the main focus of this tool is to control the event of the mouse. It is based on the SAPI and .NET technology also it integrates with the office package. It also has facility to add more commands. The feature of this tool is to minimize the mouse clicks or keyboard input.

### F.  JULIUS

This tool is a high performance in which it acts as two-pass large vocabulary continuous speech recognition which also means as LVCSR. It act as the decoder software for the speech researchers and the developers which can perform the real time decoding in the most PCs. The platform for this is tool is Linux, Windows and in the Unix. It is

free software but the main drawback is that the speech should be in continuous. There are two models in this tool they are language model and acoustic model.

### G.  SAPI

SAPI stands for Speech Application Programming Interface which is an API for the speech recognition tool. This is the product of Microsoft Corporation which has open source environment. This tool has exposed many interfaces in the form of libraries in which the functionality to bound the automation in early, in this we need to import the type of the library. SAPI can perform two functions in which the text to speech and speech to text this can be performed in the form in which option should be chosen in similar way by connecting the microphone and speaking to it and can infer that the spoken words are appear in the desktop. Before this process the Voice training should be done for the better result. The recognition is take place by means giving the input to the speech engine and by loading the dictation tool and also we need to specify the Grammar that we need to use, and then need to set the default audio input. In the code Callback function is used to perform this conversion by process all the speech events and generate the event ID to identify the speech event which has occurred.

### H.  DICTATION PRO

The Dictation pro is one of the speech recognition tool in which it is mainly used to perform the learning activity and it will convert the audio part into the text format. This tool is fast and easy to use also save time in the form of minimum keystrokes and mouse clicks. By adding the commonly used words, terminology, and technological terms to the vocabulary set of Dictation Pro. The main features of this tool are

- This tool is used for the Speech to text conversion.
- Open source tool.
- In which it will perform one of the learning activity.
- It will convert every voice into text format.

## V. VOICE RECOGNITION

The voice recognition should be done for all the speech to text tools in which it is used to analyze the modulation of the user, because people have different voice modulations. The ultimate goal of the technology is to produce a system that can recognize all words spoken by any person with 100% accuracy.  The training session makes the tool to learn which help to interpret the speech and voice which leads to increase the accuracy.

## VI. ADVANTAGES OF SPEECH TO TEXT TOOLS

1. The speaking words will be appearing on the screen no need to type.
2. All the typing and clicking process can be done just by speaking it.
3. Highly accurate based on the speech recognition tool.
4. The custom voice commands are created to insert the frequently used text.

## VII. COMPARISON OF TOOLS

Here by we have compared some tools which are to solve our problem, based upon working of these tools the comparison is performed in table1.

| TOOL | OPENSOURCE | FEATURES | TRAINING | PERFORMANCE | USAGE | ACCURACY |
|---|---|---|---|---|---|---|
| VOXSIGMA | Yes | Fair | Medium | Good | Fair | - |
| DRAGON | Yes | Good | Easy | Good | Good | 99% |
| SPEAKONIA | Yes | Fair | Medium | Good | Fair | - |
| TAZTI | Yes | Fair | Difficult | Fair | Fair | - |
| E-SPEAKING | Yes | Fair | Difficult | Fair | Fair | - |
| JULIUS | Yes | Fair | Difficult | Fair | Fair | - |
| SAPI | No | Fair | Difficult | Good | Fair | Average |
| DICTATION PRO | Yes | Excellent | Easy | Good | Good | 99% |

TABLE1  COMPARISONS OF TOOLS

## VIII. EXPERIMENTAL RESULTS

By comparing all these tools we have found that the dictation pro will be good in our concept because the functionality of this tool is best compare to other, experimentally saying we have set one large volume of audio file to this tool to annotate the most common word by means speech and the text file. Then the output will be in the form of most expected similar words in both the files and thus the annotated result was too good and most accuracy for the voice recognition process as the first step.

## IX. CONCLUSION

Thus with this experiment result we conclude that, Dictation pro tool will be suit and it is easy to work also it will produce more accuracy and ease of working environment which is an open source tool. We are going to continue our work with this tool to identify the still more semantic identification and the annotated files or words of the statement by using various functions which are all discussed in the starting of the paper. We assume that based upon the input file the functionality may be vary so need to analyze those as the future enhancement.

## REFERENCES

1.  Mark J. Weal, Danius T. Michaelides, Kevin Page, David C. De Roure, Fellow, IEEE, Eloise Monger, and Mary Gobbi "Semantic Annotation of Ubiquitous Learning Environments"
2.  www.techmatrix.org.
3.  www.literacymatters.org.
4.  Speakonia - Free download and software reviews - CNETDownload.com.http://download.cnet.com/Speakonia/3000-7239_4-10125328.html#ixzz2h205HK8P
5.  E-Speaking Voice and Speech Recognition - Free download and software reviews - CNET Download.com http://download.cnet.com/e-Speaking-Voice-and-Speech-Recognition/3000-7239_4-10371457.html#ixzz2h21mq319
6.  http://www.vocapia.com
7.  http://www.isi.edu/natural-language/mt/nitrogen