1 Group Id

Group ID 4

2 Project Title

Dictionary Application with Speech Recognition and Speech Synthesis

3 Project Option

Entrepreneurship Project

4 Internal Guide

Prof. Vijayendra Gaikwad

5 Sponsorship and External Guide

Not yet sponsored

6 Technical Keywords

1. Speech Recognition
   (a) Voice Input
   (b) ADC
   (c) Acoustic Model
   (d) Language Model
2. Searching and Sorting
3. Speech Synthesis
   (a) Formant Synthesis
   (b) Concatenation Synthesis
   (c) Articulatory Synthesis
4. Sphinx
   (a) Front End
   (b) Linguist
   (c) Decoder

7 Problem Statement

To develop an application which will take input in the form of users voice and recognize the word and finds the meaning from availble data and it will be displayed or dictated to the user using speech synthesis.

8 Abstract

We are developing a speech-to-text and text-to-speech input-output method for our Voice Based Dictionary system. The system will be implementing/using Java Sphinx-4 for Speech Synthesis and Speech Recognition. The application or the system makes use of Java Swing API for graphical user interface (GUI). The system will be deployed in the background using Java Multi-Threading support.To invoke the system the user will have to pronounce a particular word after which the application ask for the word. Once the system wakes up the immediate word listened by the system will be taken as an actual word which needs to be searched in the Database. If the similar word or the accurate word is found, then its meaning will be shown or it will be dictated in speech format.Also it can show the variations of the word.

9 Goals and Objectives

- Minimal HCI
- Offline Dictionary.
- Easy to Access
- Voice based Search
- User friendly

10 Relevant mathematics associated with the Project

System Description:

- Input Set :

  I= { W,X }
  Where,
  W=w1,w2,w3.....wn , Be the set of Word Sequence.

  X=x1,x2,x3.....xn , Be the set of acoustic observations.

- Output Set:
  O={ Wo,Mo }

  Where,
  Wo is the word matched/found.
  Mo be the set of meanings found for current word.

- Functions :

  listen() : This will take users voice as an input.

  recognize() : This function will convert the voice into particular word.

$$W = arg_w max P(W)P(X|W)$$

$$W = arg_w max \frac{P(W)P(X|W)}{P(X)}$$

$$W = arg_w max P(W|X)$$

search() : This function will search for the word and it meaning.

synthesis() : It will convert the found meaning into the verbal format.

display() : This function will show the word with its meaning.

- Success Conditions:

  User will get the words with meanings.

- Failure Conditions:

  Word is unrecognised or Word is not found.

11 Names of Conferences / Journals where papers can be published

- IEEE/ACM Conference/Journal 1
- Conferences/workshops in IITs
- Central Universities or SPPU Conferences
- IEEE/ACM Conference/Journal 2

12 Review of Conference/Journal Papers supporting Project idea

SR NO	Name Of Paper	Publisher	Authors	Description
1	Open source voice creation toolkit for the MARY TTS Platform	3 Proceedings of Interspeech 2011, Florence, Italy, ISCA, 8/2011	Marc Schrder, Marcela Charfuelan Oliva, Sathish Pammi, Ingmar Steiner	This paper describes an open source voice creation toolkit that supports the creation of unit selection and HMM-based voices, for the MARY (Modular Architecture for Research on speech Synthesis) TTS platform. The toolkit can be easily employed to create voices in the languages already supported by MARY TTS, but also provides the tools and generic reusable run-time system modules to add new languages.
2	Spoken Term Detection based on Acoustic Speech Segmentation	2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)	Nadia Benati,Halima Bahi	This paper presents a system for spoken term detection in a continuousspeech stream. Spoken terms are predefined through a set of acoustic examples provided by the users. Spoken term detection proceeds in two steps: speech segmentation and term verification. We suggest the use of an acoustic-based algorithm for the segmentation which exploits acoustic particularities of the speech stream to detect word frontiers.
3	Text to Speech Synthesizer-Formant Synthesis	2017 International Conference on Nascent Technologies in the Engineering Field (ICNTE-2017)	Sneha Lukose ,Savitha S. Upadhya	In this paper, different methods of text to speech synthesizer techniques are discussed to produce intelligible and natural output and a vowel synthesizer using cascade formant technique is implemented. A text to speech output is based on generating corresponding sound output when the text is inputted.
4	The CMU SPHINX-4 Speech Recognition System	Sun Microsystems Laboratories, USA, Carnegie Mellon University, USA, Mitsubishi Electric Research Labs, USA, University of California, Santa Cruz, USA	Paul Lamere, Philip Kwok, Evandro Gouvła, Bhiksha Raj,Rita Singh, William Walker, Manfred Warmuth, Peter Wolf	The Sphinx-4 speech recognition system is the latest addition to Carnegie Mellon University's repository of Sphinx speech recognition systems. Algorithmic innovations included in the system design enable it to incorporate multiple information sources in an elegant manner. The system is entirely developed on the Java platform and is highly portable, flexible, and easier to use with multithreading. This paper describes the salient features of the Sphinx-4 decoder and includes preliminary performance measures relating to speed and accuracy.

13 Block Diagram

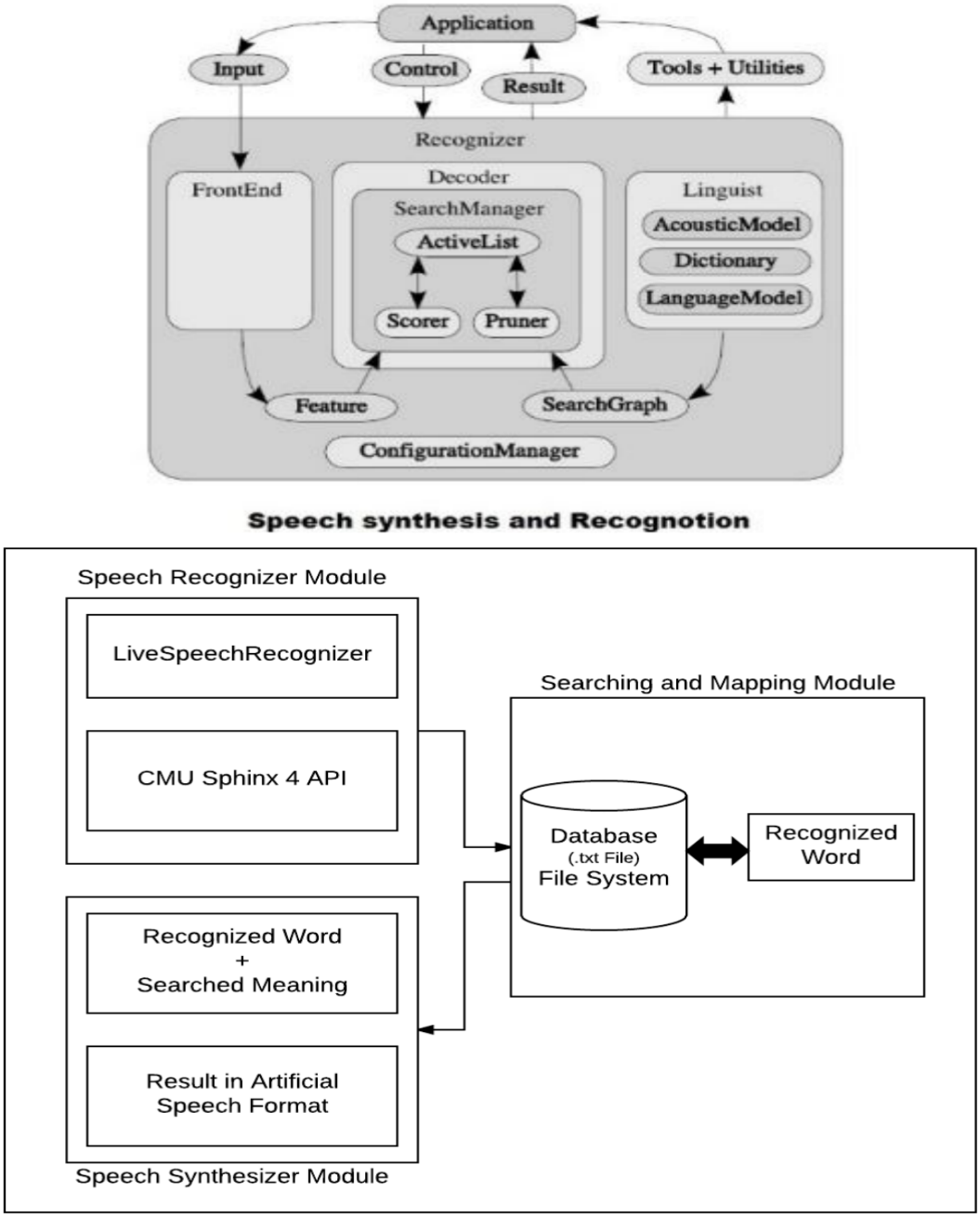
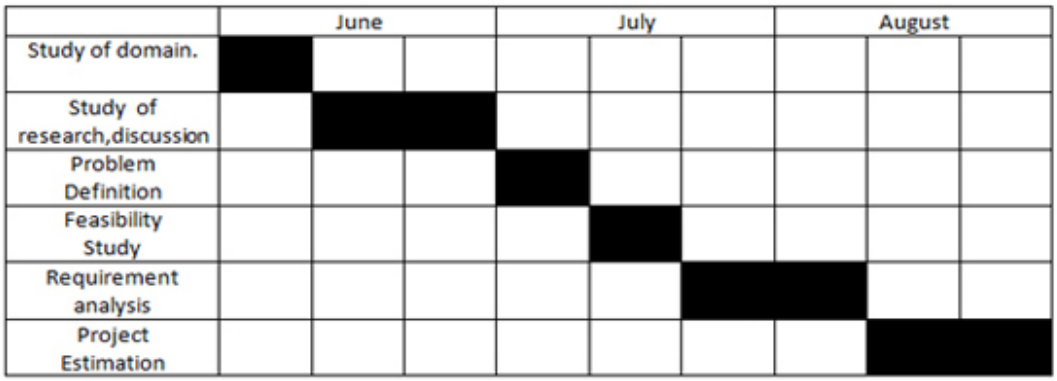
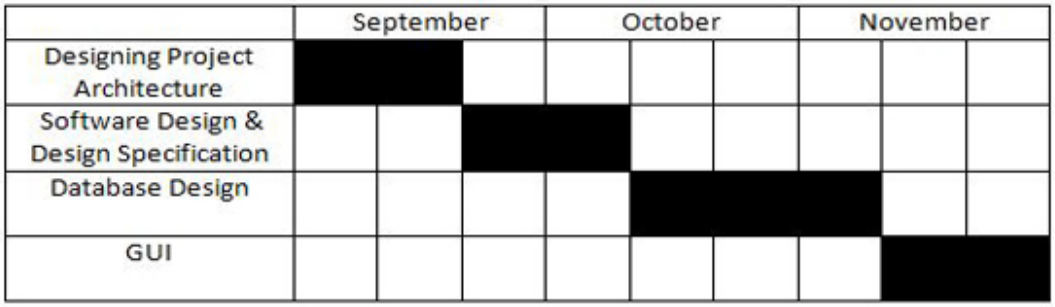


Figure 1: System Architecture

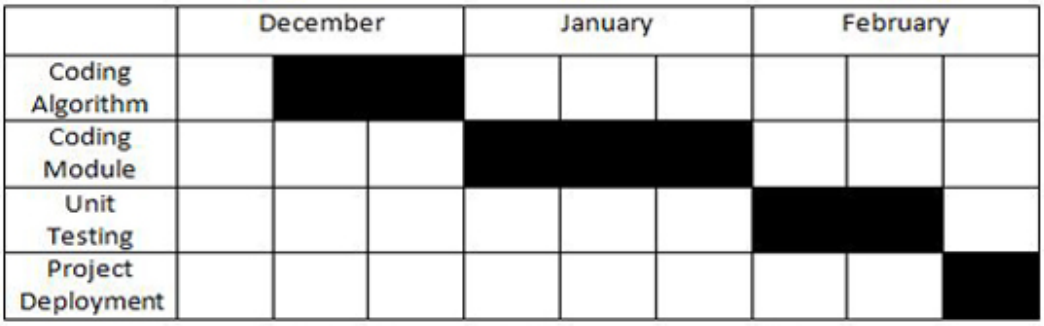
14 Project Estimated Schedule



Analytical Phase



Design Phase



Coding and Deployment Phase

Prof.Vijayendra Gaikwad
Internal Guide

Prof. Rama Gaikwad
Project Coordinator

Prof. Manoj Mulik
HOD
Dept. of Computer Engineering