



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 6      Issue: IV      Month of publication: April 2018**

**DOI: <http://doi.org/10.22214/ijraset.2018.4331>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Automatic Speech Recognition: - Ease of Typing

Nakshatra Sain<sup>1</sup>, Sagar Kapse<sup>2</sup>, Malika Bindal<sup>3</sup>, Surbhi<sup>4</sup>, Prof. Prashant S. Chavan<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup>Department of Information Technology, Bharati Vidyapeeth (To Be Deemed University) College of Engineering Pune

**Abstract:** A Web Based Artificial Speech Recognition Tool Project to eliminate typing and implement ease of use. This system is all about the online web based speech recognition tool this means that user just have to talk or speak some sentence or word into the system and then system will fetch the word or sentence from the database and displays the sentences or word which has speak by the user. In the existing online system user can face so much problems because of typing the text into the system.

They have to waste so much time for making reports, documents, also for searching information on online and also to find any document in the computer system. Speech Recognition system requires 'training' of corpus for a smooth flow of system. This tool will be helpful in providing data for training.

**Keywords:** Automatic Speech Recognition, Hidden Markov, Acoustic model, MFCC, ASR

## I. INTRODUCTION

A speech corpus is a collection of audio recordings of spoken language together with their respective annotations and Meta information. Essentially the discourse corpus is the accumulation of discourse signs and its explanation, metadata, and reports, is the reason for both breaking down the qualities of discourse flags and creating discourse amalgamation and acknowledgment frameworks. It isn't just for TTS innovation, yet in addition utilized for ASR (Programmed Discourse Acknowledgment) and phonetic research, is discourse corpus critical. For phonetic research, discourse corpus can give differing and precise information to enable specialists to discover the tenets of dialects. Also, for ASR, keeping in mind the end goal to "prepare" the framework to "see" any of the speakers' voices, a discourse corpus with an extraordinary limit is essential. The principle preferred standpoint of the measurable information of a discourse corpus, the ASR framework can change discourse signals into content strings by utilizing phonological, semantic, and stochastic investigation This is the basic idea behind why ASR can "understand" human's voice. Basically there are two types of speech corpus. Read Speech - which includes: Book excerpts, Broadcast news, Lists of words and Spontaneous Speech which includes Dialogs and narratives

## II. HISTORY

Discourse Acknowledgment has been continuous for over 80 years. Over that period there have been no less than 4 age of methodologies, and a fifth era is being figured in light of ebb and flow explore subjects. In 2001, PC discourse acknowledgment had achieved 80% exactness and no further advance was accounted for till 2010. In 2010, Google included "customized acknowledgment" to Voice Hunt on Android telephones, with the goal that the product could record clients' voice pursuits and deliver a more exact discourse demonstrate.

The organization likewise added Voice Inquiry to its Chrome Program in mid-2011. Like Google's Voice Pursuit, Siri depends on cloud-based preparing. It draws on its information about the speaker to produce a relevant answer and reacts to voice enter. Parallel preparing strategies utilizing mixes of Well and acoustic-phonetic ways to deal with recognize and rectify semantic inconsistencies are utilized to expand acknowledgment choice unwavering quality and increment strength for acknowledgment of discourse in boisterous condition.

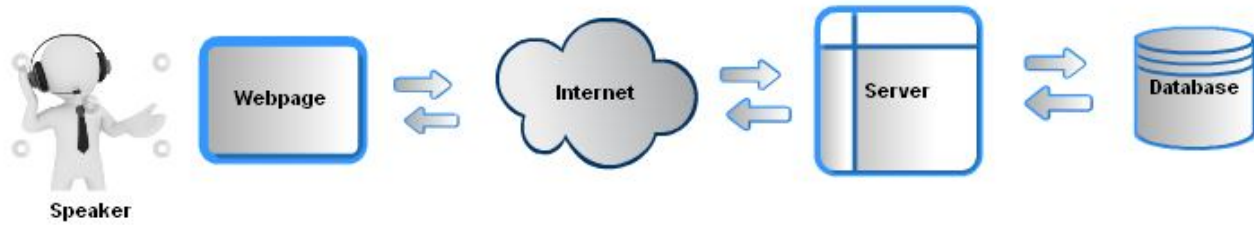
## III.SPEAKER PROFILES

The speaker data was established during the preparation of a recording session for transcription. A given set of individual recordings containing information about the speaker was listened to and the data was entered into the database containing the speaker Meta data, e.g. speaker gender, accent, age, etc.

The characteristic features of each speaker should be collected in the speaker profile set of the corpus.

For each speaker summarizing the features per speaker for example Speaker ID, Sex, Date of birth , Corpus Language , Mother tongue, Place of elementary school , Dialect region , Dialect , Level of education , Level of Computer , Proficiency ,Profession, Smoker/nonsmoker.

#### IV. SYSTEM ARCHITECTURE



The client-server architecture is used for collecting speech databases. The central server is responsible for the data storage, and it offers access to its data in the form of services. Recording clients (speaker) request prompt items from the server and return the recorded signal data to the server.

#### V. SPEECH RECOGNITION ARCHITECTURE

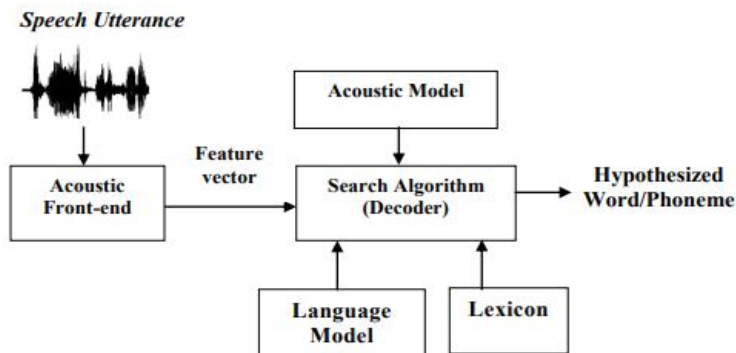


Figure 1. Speech Recognition Architecture

The speech recognition architecture is built with the help of various major components such as acoustic front end, acoustic model, lexicon, language model and decoder. While recording phase the end user will read all the words which are there in the sentences of the current sets of that application. Then recording will be saved in the database and then later the sentences which we have recorded from the user are compared with stored sentences which were there in the database. And best match is selected. This is called as pattern recognition. There are four basic components are as follows:-

##### A. Acoustic front end

Acoustic front end is nothing but the signal processing and it also extract features. The main goal of feature extraction step is that to compute a sequence of feature vectors providing with a compact representation of the given input signal. It also takes care about of converting the speech signal into its relevant features which will provides us useful information for recognition.

##### B. Acoustic model

In speech recognition architecture acoustic model is the most important knowledge source for automatic speech recognition which provides us the features for phonetic units to be recognized.

##### C. Language model

It is the collection of constraints on which the sequence of the word is accepted in the form of given language.

##### D. Decoder

The word decoder is nothing but the task to find out the most likely word in the given sequence W which was given in the observation sequence O and also in the acoustic phonetic language model. There were various problems while decoding which can be get solved through by using dynamic programming algorithms'.



## VI. CONCLUSIONS

As we come to the end the final output of our project would be a very useful System for students, faculty and everyone. The product we will create would not only be useful to a faculty, parents but also to a student.

This Web Based Artificial Speech Recognition Tool project intentions was is to computerized the tradition way of recognizing the word or the sentence without typing any single word on the system.

This system is all about the online web based speech recognition tool this means that user just have to talk or speak some sentence or word into the system and then system will fetch the word or sentence from the database and displays the sentences or word which has speak by the user.

## REFERENCES

- [1] M.A.Anusuya and S.K.Katti, "Speech Recognition by Machine: A Review", (IJCSIS) International Journal of Computer Science and Information Security, vol. 6, no. 3, pp. 181-205, 2009.
- [2] Mohit Dua, R.K.Aggarwal, Virender Kadyan and Shelza Dua, "Punjabi Automatic Speech Recognition Using HTK", IJCSI International Journal of Computer Science Issues, vol. 9, issue 4, no. 1, July 2012.
- [3] Rajesh Kumar Aggarwal and M. Dave, "Acoustic modelling problem for automatic speech recognition system: advances and refinements Part (Part II)", Int J Speech Technol, pp. 309- 320, 2011.
- [4] Kuldeep Kumar, Ankita Jain and R.K. Aggarwal, "A Hindi speech recognition system for connected words using HTK", Int. J. Computational Systems Engineering, vol. 1, no. 1, pp. 25-32, 2012.
- [5] Kuldeep Kumar R. K. Aggarwal, "Hindi speech recognition system using HTK", International Journal of Computing and Business Research, vol. 2, issue 2, May 2011



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)