



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 8      Issue: VIII      Month of publication: August 2020**

**DOI: <https://doi.org/10.22214/ijraset.2020.30985>**

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

**Call:  08813907089**

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

# Audio Word Processor

Mr. Aryan Singhal<sup>1</sup>, Mr. Archit Jain<sup>2</sup>, Mr. Anuj Gupta<sup>3</sup>, Mr. Abhishek Kumar<sup>4</sup>, Mr. Ashwani Tyagi<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup>CSE Students, Computer Science and Engineering Department, IMS Engineering College, Ghaziabad, U.P., India

## I. INTRODUCTION

Dimorphism is the property of voice that is highly observed in human beings. Intonation, speech rate, and duration are certain characteristics that distinguish human voices. The sound of each individual's voice is entirely unique not only because of the actual shape and size of an individual's vocal cords but also due to the size and shape of the rest of that person's body, especially the vocal tract, and the manner in which the speech sounds are habitually formed and articulated. (It is this latter aspect of the sound of the voice that can be mimicked by skilled performers.)

There are 3 main characteristics of sound: frequency, amplitude, and waveform. In human terms: pitch, volume and timbre. Two people can sing the same note (pitch) at exactly the same loudness (amplitude) and still you can tell the difference between both voices. That's because of their personal timbre. The timbre builds up from many factors, among them the most important are: The vocal cords and the resonators (Note: you can use "vocal cords" and "vocal folds" interchangeably.).

The vocal cords are a couple of opposing folds in the larynx that vibrate in the same way as when you force air through your lips. The cords can be longer or shorter, and thinner or thicker. As air passes through the stretched cords they vibrate in a particular way according to their physical characteristics and that germinal sound changes according to the shape of the vocal tract. That's why exact replicate can't be produced but similar voice can be produced.

## II. OBJECTIVE

To convert human audio into text and whatever changes made to text are reflected back to human audio (with similar human voice). Edit audio by editing text. Correct your voice recordings by simply typing.

## III. CHALLENGES

- 1) *Data Preprocessing*: While fetching data is a challenge in itself, the data we have gathered may have several impurities that we may want to remove. Various giniimpurities and NAN values are needed to be get rid of. We can use fillna method from pandas library in python to fill all unavailable values with some agreed value, That value may be mean or median of whole column.
- 2) *Expertise in ML*: We are not ashamed to say that we have 0-year experience in the field of machine learning. We are fresher so we cannot work like professionals. While developing such huge system it is a challenge whether we as a fresher can deal with various problems encountered in the project.
- 3) *Evaluating Characteristics*: There are 3 main characteristics of sound: frequency, amplitude, and waveform. In human terms: pitch, volume and timbre.
- 4) *Accuracy*: Biggest factor is the accuracy of the system. But with the quality of data it is sure that you can not increase the accuracy after some specified threshold. So we need to work continuously on the data to make it sure that our system is predicting with good accuracy.

### A. Hardware Requirement

- 1) i3 Processor Based Computer
- 2) 1GB-Ram
- 3) 5 GB Hard Disk
- 4) Internet Connection

### B. Software Requirement

- 1) Windows 7 or higher
- 2) Microphone and speaker
- 3) Web Browser



### C. Technology

Python and JavaScript is used to program. HTML and CSS is used for front end. Machine learning is used to get the results based on symptoms.

### D. About the Language

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. Python features a dynamic type system and automatic memory management . Some of the major changes included for Python 3.0 were: Changing print so that it is a built-in function, not a statement. This made it easier to change a module to use a different print function, as well as making the syntax more regular. In Python 2.6 and 2.7 print() is available as a builtin but is masked by the print statement syntax, which can be disabled by entering from `__future__ import print_function` at the top of the file. Removal of the Python 2 input function, and the renaming of the raw\_input function to input. Python 3's input function behaves like Python 2's raw\_input function, in that the input is always returned as a string rather than being evaluated as an expression. JavaScript (often shortened to JS) is a lightweight, interpreted, object-oriented language with first-class functions, and is best known as the scripting language for Web pages, but it's used in many non-browser environments as well. It is a prototype-based, multi-paradigm scripting language that is dynamic, and supports object-oriented, imperative, and functional programming styles. JavaScript runs on the client side of the web, which can be used to design / program how the web pages behave on the occurrence of an event. JavaScript is an easy to learn and also powerful scripting language, widely used for controlling web page behavior.

## IV. CONCLUSION

Speech to text synthesis is a rapidly growing aspect of computer technology and is increasingly playing a more important role in the way we interact with the system and interfaces across a variety of platforms. We have identified the various operations and processes involved in the text to speech synthesis. We have also developed a very simple and attractive graphical user interface that allows the user to type in his/her text provided in the text field in the application.

Our system has interfaces with a speech to text and text editor through which we can convert the edited text into speech. Another area of further work is the implementation of a text to speech and speech to text system on other platforms, such as telephony systems, ATM machines, video games and any other platforms where the text to speech technology would be an added advantage and increase functionality. This also provides benefits to the visual aid people.

## REFERENCE

- [1] <https://www.descript.com/>
- [2] <https://economictimes.indiatimes.com/tech/software/new-software-can-edit-human-voice-like-text/articleshow/58701058.cms?from=mdr>
- [3] <https://www.sciencedaily.com/releases/2017/05/170515091121.htm>





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)