Text – To – Audio Converter (TTA)

Swati Kadlag[1], Ankit Purohit[2], Abhishek Kumar[3], Shraddha Mhaske[4]

[1]Department of E&TC Engg, Symbiosis Institute of Technology, Pune, India
(Swati Kadlag, swatik@sitpune.edu.in)

[2]Symbiosis Institute of Technology, Pune, India
(AnkitPurohit, ankit.purohit@sitpune.edu.in)

[3]Symbiosis Institute of Technology, Pune, India
(Abhishek Kumar, abhishek.mahto@sitpune.edu.in)

[4]Symbiosis Institute of Technology, Pune, India
(ShraddhaMhaske, mhaske.shraddha@sitpune.edu.in)

[5]Symbiosis Institute of Technology, Pune, India
(SamyakVaidya, samyak.vaidya@sitpune.edu.in)

Abstract

This paper aims to give an overview of Text- to- Speech converter using JAVA script. This report presents the development of a progressive web application of Text- to- Audio converter which is capable of performing different tasks. It can convert the text which is typed in the text box into audio and can also convert a pdf text to audio.
Text-to-Audio Converter is a technology that converts written text from a descriptive medium to a spoken language, allowing the end user to easily understand what is written via the audio output.

Keywords: Speech Synthesis, FreeTTS, JAVA Script, Speech, Text-To-Audio, Phonetic Analysis.

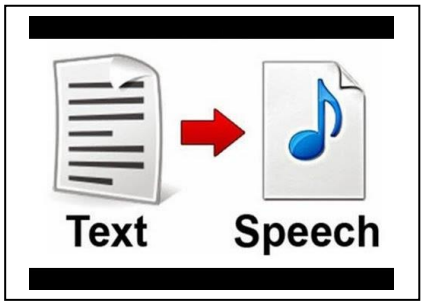

Fig. 1

Introduction

Speech synthesis is the process of synthesising human speech artificially. A technology used for this purpose is known as speech-synthesiser, and implementation of speech-synthesiser can be done through hardware and software.A (T.T.S) system converts a text into speech; other systems convert symbolic linguistic representations such as phonetic transcriptions into speech. (T.T.S) is a linguistic or text-to-speech conversion convention. It's now used everywhere in reading audio systems for the visually impaired. However, in recent years, technology for (T.T.S) conversion has expanded far beyond the disabled community, becoming a significant complement to the rapidly increasing use of digital voice storage for voice mail and VOIP.(T.T.S) encourages people to enjoy and also gives them the option of consuming content on the go, moving content away from the computer screen and into whatever environment is most convenient for them. (T.T.S) can also be a very useful tool for people with visual impairment. Reading a large amount of content on a small screen is not always easy for those who access content through mobile devices.

Background

People with literacy problems and those learning a new language often become frustrated while trying to navigate the internet due to the overwhelming amount of text. Even if they can read content with a simple understanding, many people struggle to read fluently in a second language.Dyslexics and other people who havedifficulties in learning,who cannot read text of large volumesor some different issues, faced a huge problem. But TTS made it easier to understand what is written without reading the text.

1. Objectives of the Study

The main motive of this project is to create a progressive web application for text-to-speech converter. This software will helps to bridge the gap and reduce any scope for communication gap for people having different reading

skills. It will help those people who have difficulty to read text for various reasons, an audio output is more handy to consume.

The desire to complete a demanding project in a fascinating field of study was the driving force behind this project. It was appealing to hear about a new field of computing that was not discussed in lectures.

3. Scope of the Study

1. Audio quality is more convenient to consume for people who have trouble reading text for different reasons.
2. Since not everybody has the same level of reading ability, audio output will help, bridge the chasm and eliminate the potential for miscommunication.
3. It helps the user to multitask when listening to audio.
4. Read any text on the computer:
- Web pages
- Word documents
- E-mails
- News articles
- Online Books

5. The text-to-speech reading can be captured and saved as a WAV or MP3 file, allowing students to listen to it on their MP3 or CD players later.
6. Read any text in the voice or accent of your choice (male, female, British English).The listener selects the language (e.g., American English, etc.).
7. Accuracy will continue to increase.
8. This technology can be used in a number of languages.
9. Both the public and private sectors of the business world benefit from this technology.
10. Both children and adults can comfortably use the technology.
11. For the past ten years, the aim has been to improve Text to Speech technology. Software now has a more human-like feel to it.

Literature Survey

In the late 1950s, the first speech-synthesis computer based systems were developed. In 1968, Noriko Umeda and colleagues developed the first text to speech device at Japan's Electro technical Laboratory.

5. Text – To - Speech Synthesis Defined

By definition, an audio synthesis system is one that generates synthetic audio. It is self-evident that this requires some form of feedback. What isn't obvious is what kind of input this is. The system is known as a text-to-audio (TTA) system

if the input is plain text with no additional phonetic and/or phonological information.The synthesis process begins with text input. Nowadays, this may be plain text or text that has been marked up in any way, such as HTML or something similar to JSML (Java Synthesis Mark-up Language).

6. METHODOLOGY OF IMPLEMENTATION

To use Java script to include a text to speech feature to our website, we are going to use the Web Speech API, which is going to be used for synthesis speech (converting T.T.S) and identify speech (converting speech to text).

We are going to use the following interfaces/properties:

1. SpeechSynthesis
2. SpeechSynthesisUtterance
3. window SpeechSynthesis

JAVASCRIPT SPEECHSYNTHESIS INTERFACE

The speech synthesis service's primary interface, which manage the synthesis or production of the speech based on text input. This interface is used for accessing the device's supported voices as well as start, stop, pause, and resume speech.

The techniques available in this Interface are as follows:

Speak(): is a function that allows you to speak. Add the utterance(object of SpeechSynthesisUtterance) to the queue, and it will be spoken if no other utterances are waiting to be spoken. This is the function we'll be employing as well.

Pause(): pauses the currently playing audio.

Resume(): This function is used to resume a stopped speech.

Cancel(): Cancels all pending utterances or speech that hasn’t been played yet.

getVoices(): returns a list of all voices that are supported for the device.

JAVASCRIPTSPEECHSYNTHESISUTTERANCEINTERFACE

In this interface we actually make the speech or utterance from the text given, including language, volume, voice, speech rate, and so on. We give the object to the Speech Synthesis object's speak() method to play the speech after we've developed it for this interface.

Lang: To decide and set the speech language.

Pitch: The voice pitch that will deliver the utterance must be calculated and set.

Rate: Obtaining and changing the rate at which the utterance will be transmitted is referred to as rate.

Volume: Obtaining and changing the volume is what volume is all about.

Text: To retrieve and setting the text to be read aloud.

Voice :To obtain or configure the voice that will be used.

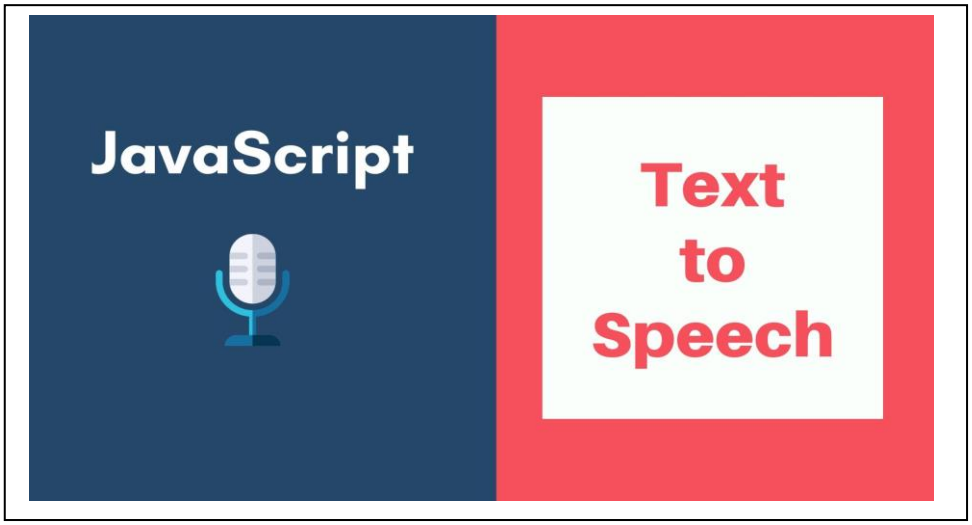

Fig.2

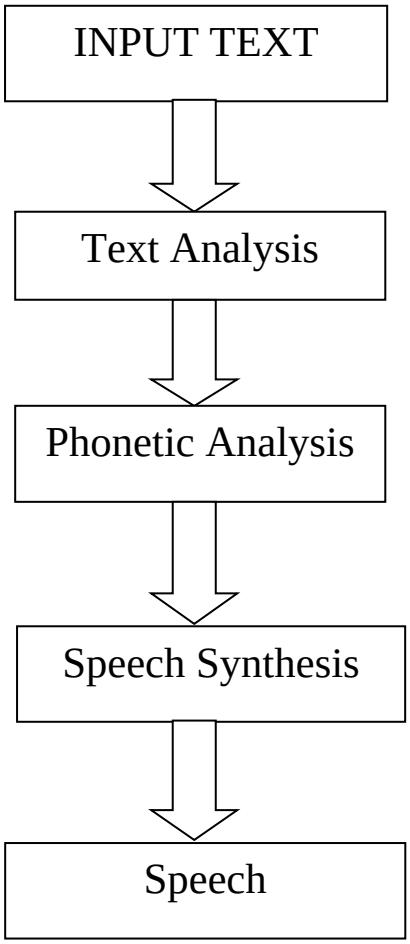

Fig. 3 Block Diagram of Text- to- Speech Synthesis

Software Architecture:

- Software Type: Web Application
- Frontend: php, Javascript, html
- Database: mysql
- Web URL: localhost/tos

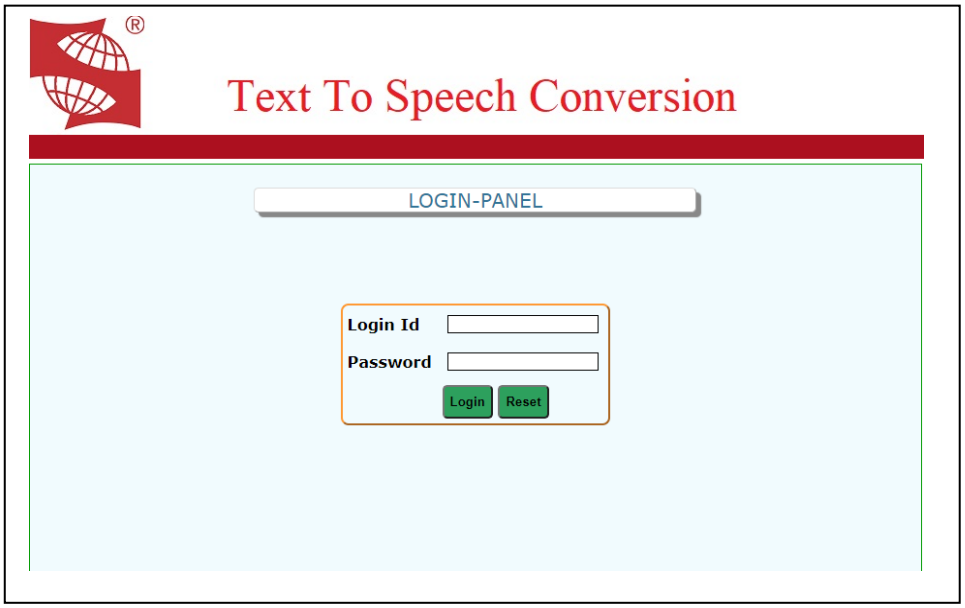

Fig. 4 Login page

Fig. 5 Home page

There will be two ways by which user can convert text-to-speech in our progressive web application

1) User can directly enter the text in the text box
2) User can directly upload a pdf

9. Conclusion and Recommendation

An cost-effective, remotely controllable using internet with easy user friendly interface, Text- to- Audio converter has been realised which was discussed in this report. This technology will be very useful for the people who are facing reading issues and this technology will surely contribute for the betterment of the society.

REFERENCES

[1] Lemmetty, S., 1999. Review of Speech Syn1thesis Technology. Masters Dissertation,

Helsinki University Of Technology.

[2] Dutoit, T., 1993. High quality text-to-speech synthesis of the French language.Doctoral

dissertation, FaculтePolytechnique de Mons.

[3] Suendermann, D., Höge, H., and Black, A., 2010. Challenges in Speech Synthesis. Chen, F.,

Jokinen, K.,(eds.), Speech Technology, Springer Science + Business Media LLC.

[4] Allen, J., Hunnicutt, M. S., Klatt D., 1987. From Text to Speech: The MITalk system.

Cambridge University Press.

SpeechSynthesis.Springer.

[7] van Santen, J.P.H., 1994. Assignment of segmental duration in text-to-speech synthesis.

Computer Speech & Language, Volume 8, Issue 2, Pages 95–128

[8] Wasala, A., Weerasinghe R. , and Gamage, K., 2006, Sinhala Grapheme-to-Phoneme

Conversion and Rules for Schwaepenthesis. Proceedings of the COLING/ACL 2006 Main

Conference Poster Sessions, Sydney, Australia, pp. 890-897.

[9] Text-to-speech (TTS) Overview. In Voice RSS Website. Retrieved February 21, 2014, fromhttp://www.voicerss.org/tts/

[10] Text-to-speech technology: In Linguatec Language Technology Website. RetrievedFebruary21,2014,Fromhttp://www.linguatec.net/products/tts/information/technology

http://mediawiremobile.com/news/text-to-speech-applications-benefits-and-uses/#:~:text=TTS%20technology%20allows%20them%20to,that's%20convenient%20for%20the%20consumer.

https://www.ijert.org/research/implementation-of-text-to-speech-conversion-IJERTV3IS030548.pdf