A
Project Report
On

Speech-synthesizer

Submitted in partial fulfillment of the requirement for the award of
BACHELOR OF TECHNOLOGY

in

Computer Science and Engineering

UNDER THE GUIDANCE OF

Mr. Devendra Kumar Singh
Assistant Professor

SUBMITTED BY

Shreyas Kumar Thakur	**Santosh Yadav**
Roll no. 18103053	**Roll no. 18103051**

GURU GHASIDAS VISHWAVIDYALAYA, BILASPUR, INDIA

February,2021

INSTITUTE OF TECHNOLOGY

GURU GHASIDAS VISHWAVIDYALAYA ,BILASPUR (C.G.)

CERTIFICATE

I hereby certify that the work which is being presented in the B.Tech. Minor Project Report entitled "**Speech-synthesizer**", in partial fulfillment of the requirements for the award of the Bachelor of Technology in Computer Science and Engineering and submitted to the Department of Computer Science and Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya(A Central University), Bilaspur, Chhattisgarh, India is an authentic record of my own work carried out during a period from July 2020 to December 2020(5th semester) under the supervision of **Mr. Devendra Kumar Singh**, Assistant Professor, CSE Department.

The matter presented in this Project Report has not been submitted by me or by anyone else for the award of any other degree elsewhere.

Signature of Student (S)

Student Name   Student Name

Shreyas Kumar Thakur(18103053)   **Santosh Yadav(18103051)**

This is to certify that the above statement made by the student(s) is correct to the best of my knowledge.

Signature of Supervisor(s)

Mr. Devendra Kumar Singh

Date: Assistant Professor

Head

Deptt. of Computer Science and Engineering

DECLARATION

We here by declare that the work presented in this dissertation entitled **"Speech-synthesizer"** submitted to the **Department of Computer Science & Engineering** , under the guidance of **Mr. Devendra Kumar Singh** has been done by us, and this dissertation embodies our own work. The work is original as it has not been earlier submitted either in part or full for any purpose before by us or anyone else.

Shreyas Kumar Thakur   **Santosh Yadav**

rollno. 18103053   rollno. 18103051

ACKNOWLEDGEMENTS

.

We would like to place on record my deep sense of gratitude to **Mr. Devendra Kumar Singh**, Assistant Professor, Dept. of Computer Science and Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya, Central University, Bilaspur, Chhattisgarh, India, for his stimulating guidance, continuous encouragement and supervision throughout the course of present work.

We express my sincere gratitude to **Dr. Alok Kumar Singh Kushwaha**, Associate Professor and HOD-Dept. of Computer Science and Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya, Central University, Bilaspur, Chhattisgarh, India for his generous guidance, help and useful suggestions.

We are extremely thankful to Prof T.V.Arjunan , Dean, School of Studies(Engg. & Tech.), Guru Ghasidas Vishwavidyalaya, Central University, Bilaspur, Chhattisgarh, India for providing us infrastructural facilities to work in, without which this work would not have been possible.

Date: Name of Students

                                                                               Shreyas Kumar Thakur(18103053)

                                                                               Santosh Yadav(18103053)

ABSTRACT

Speech synthesis (TTS) is defined as the artificial production of human voices. The main use (and what induced its creation) is the ability to translate a text into spoken speech automatically.

Unlike speech recognition systems that use phonemes (the smallest units of sound) in the first place to cut out sentences, TTS will be based on what are known as graphemes: the letters and groups of letters that transcribe a phoneme. This means that the basic resource is not the sound, but the text. This is usually done in two steps. The first will cut the text into sentences and words (our famous graphemes) and assign phonetic transcriptions, the pronunciation, to all these groups. Once the different text/phonetic groups have been identified, the second step consists of converting these linguistic representations into sound. In other words, to read these indications to produce a voice that will read the information.

INTRODUCTION

Need of the Software:

Text-to-speech (TTS) is a very popular assistive technology in which a computer or tablet reads the words on the screen out loud to the user. This technology is popular among students who have difficulties with reading, especially those who struggle with decoding. By presenting the words auditorily, the student can focus on the meaning of words instead of spending all their brain power trying to sound out the words. While this can help students work around their reading difficulties and access the classroom material, this technology does not assist students in developing reading skills.

Text-to-Speech

Text-to-Speech (TTS) is a useful technology that converts any text into a speech signal.It can be utilized for various purposes, e.g. car navigation, announcements in railway stations, response services in telecommunications, and e-mail reading. Text-to-speech (TTS) is a type of speech synthesis application that is used to create a spoken sound version of the text in a computer document, such as a help file or a Web page. TTS can enable the reading of computer display information for the visually challenged person, or may simply be used to augment the reading of a text message. Current TTS applications include voice-enabled e-mail and spoken prompts in voice response systems. TTS is often used with voice recognition programs.

Speech Synthesizer

Text → Natural Language Processing → Digital Signal Processing → Speech
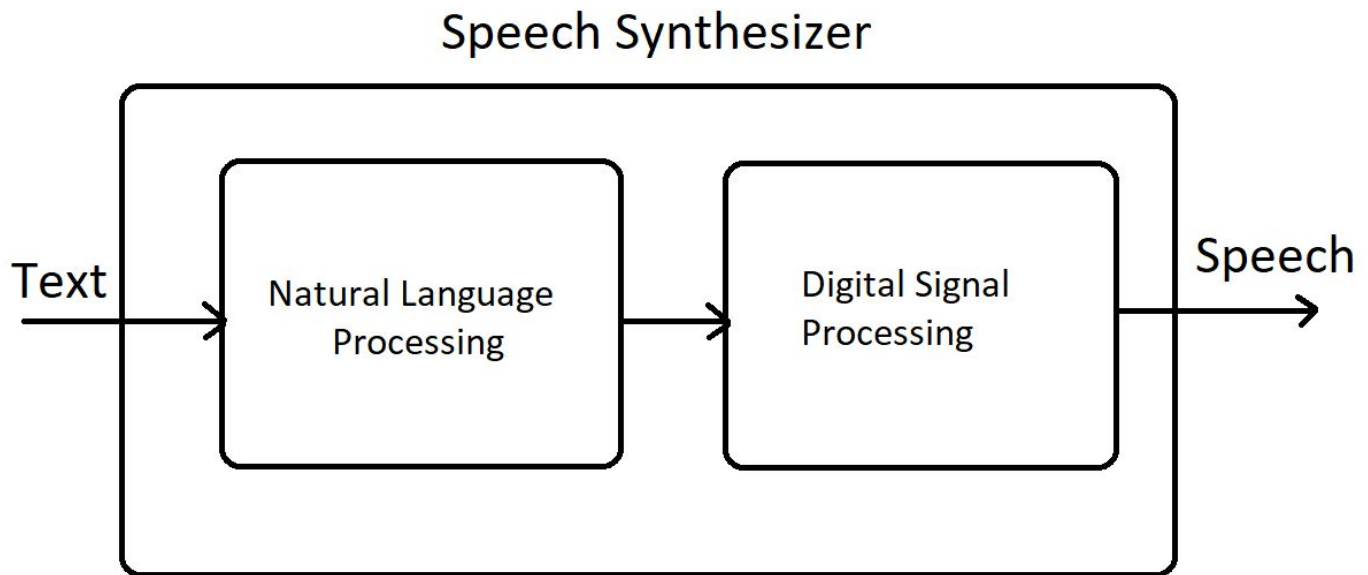
Audiobooks

An audiobook is a recording or voiceover of a book or other work read aloud. In our project , we are making an audiobook which takes a PDF file as an Input and returns audio file as Output to the user. Audiobooks have been used to teach children to read and to increase reading comprehension. They are also useful for the blind. The National Library of Congress in the U.S. and the CNIB Library in Canada provide free audiobook library services to the visually impaired; requested books are mailed out (at no cost) to clients.

Spell-Checker

In software, a spell checker (or spell check) is a software feature that checks for misspellings in a text. Spell-checking features are often embedded in software or services, such as a word processor, email client, electronic dictionary, or search engine.

Libraries and Tools used :

GTTS library :

gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate's text-to-speech API. Write spoken mp3 data to a file, a file-like object (bytestring) for further audio manipulation, or stdout. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

'Playsound' library :

To play the processed audio file after text-to-speech conversion.

'PyPDF2' library :

A Pure-Python library built as a PDF toolkit. It is capable of extracting document information (title, author, …). By being Pure-Python, it should run on any Python platform without any dependencies on external libraries. It can also work entirely on StringIO objects rather than file streams, allowing for PDF manipulation in memory. It is therefore a useful tool for websites that manage or manipulate PDFs.

'Tkinter' library :

This package ("Tk interface") is the standard Python interface to the Tk GUI toolkit.Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit, and is Python's de facto standard GUI. Tkinter is included with standard Linux, Microsoft Windows and Mac OS X installs of Python. The name Tkinter comes from the Tk interface.

'Textblob' library :

TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP)tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.

Language used : Python

Python is a high-level programming language that has English-like syntax. This makes it easier to read and understand the code.
Python is really easy to pick up and learn, that is why a lot of people recommend Python to beginners.You need less lines of code to perform the same task as compared to other major languages like C/C++ and Java. Advantages of python :

-Versatile, Easy to Use and Fast to Develop.

-Open Source with a Vibrant Community.

-Has All the Libraries You Can Imagine.

-Great for Prototypes - You Can Do More with Less Code.

Text-editor used : VS Code

Visual Studio Code is a lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux. It comes with built-in support for JavaScript, TypeScript and Node.js and has a rich ecosystem of extensions for other languages (such as C++, C#, Java, Python, PHP, Go) and runtimes (such as .NET and Unity).

In This Project, we have combined three different modules to work together. These modules are text-to-speech module, pdf-to-speech(audiobook) module and Spell-checker module.

Project Output :
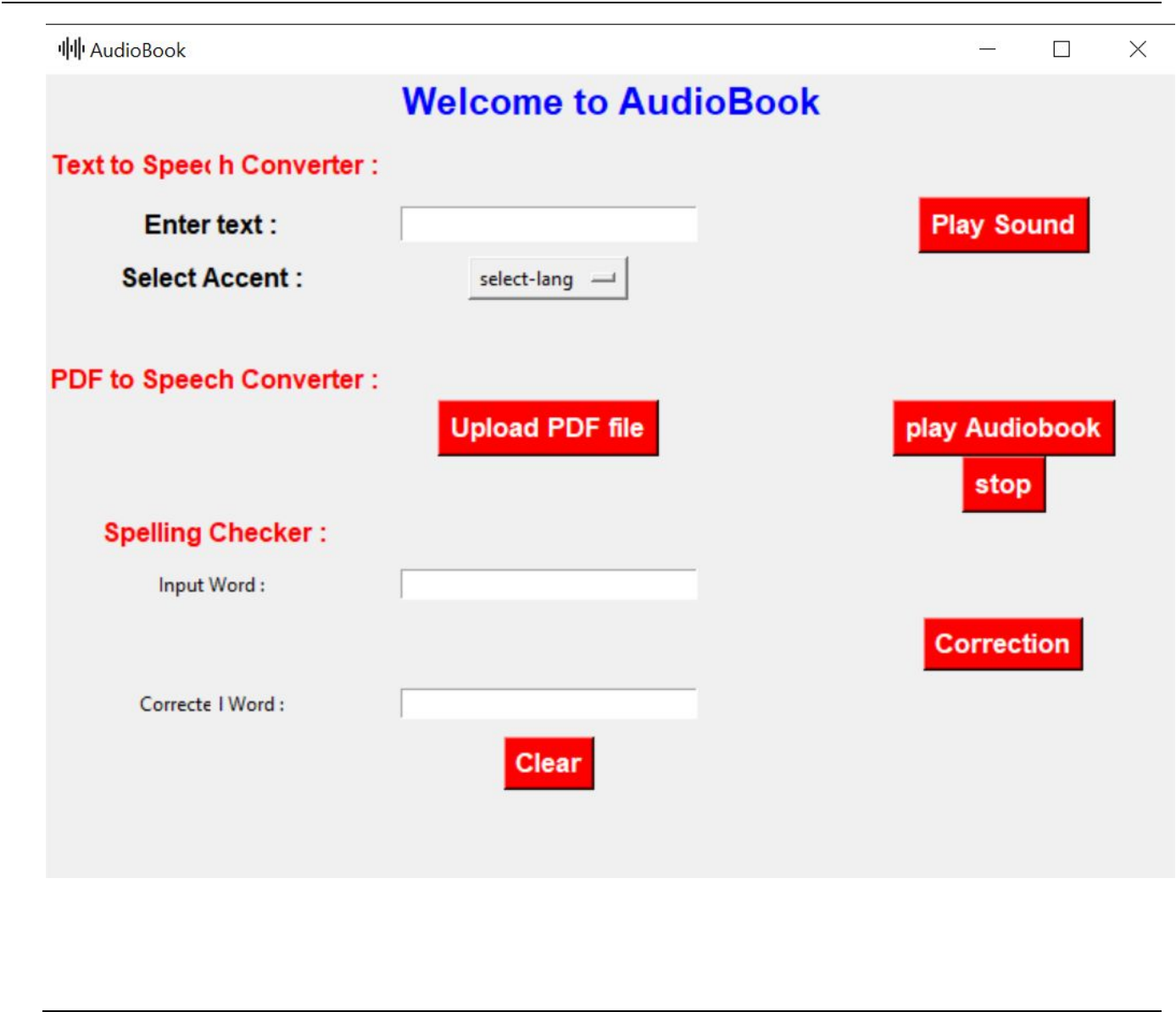
How it Works ?

Text-to-speech module :

In TTS module, the GUI has a textbox with a label which says "Enter text". When the user enter the text in the textbox and clicks the "Play Sound" button, the text data is sent to the gtts API for the conversion. As gtts API works online, it highly depends on the speed of internet connection. The gtts library sends the data to google servers and fetches the processed audio information. This audio is then saved in mp3 format in local computer. With the help of "Playsound" library, we then play the audio file.

As this module is based on Google's API , it gives the liberty to add further datasets to the predefined datasets, so that users can modify the outputs for desirable accents and pronunciation.

PDF-Reader Module :

In this module, User needs to upload the PDF file to be read out aloud. When the PDF gets uploaded, the PyPDF2 library reads the PDF, and counts the number of pages. Using loop, we extract the texts from each page of the file. Now, as we want to continuously read the texts without any gap between pages, we will append the texts from all the pages in continuous manner using "append" function. Now that we have extracted all the text, we will pass it to the gtts api. The gtts library sends the data to google servers and fetches the processed audio

information. This audio is then saved in mp3 format in local computer. With the help of "Playsound" library, we then play the audio file.

Spell-checker Module :

In this module, the user needs to enter the sentence which is to be corrected. After getting the text, we pass it to the Textblob library. This library has various uses like part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation. With Textblob library, we will make correction to
the   sentence using "correct()" function. Textblob returns the correct statement. We then print this output in the textbox.

The accuracy of textblob is limited to 60%, hence it sometimes returns wrong corrections. However, it is possible to train the textblob for better accuracy , by adding more precise datasets.

Conclusion :

Text-to-speech is a great help to the visually impaired people or people with other disabilities as it can help them by assisting in the text to speech translation. There are also many ideas possible with the gTTS module and it can be used for other languages as well. Similarly, both pdf-reader and spell-checker are helpful in saving time. One can use pdf-reader while doing household chores saving both time and money.