International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 02 | Feb 2024 www.irjet.net p-ISSN: 2395-0072

Real Time Text to Speech Converter and Translator

Prof. Heena Patil[1], Prof. Archana Sahu[2], Yash Patil [3], Aditya Kesari[4], Sameer Jadhav[5], Masir Momin[6]

[1]Head of Department, Dept. of AIML Diploma, ARMIET, Maharashtra, India
[2]Lecturer, Dept. of AIML Diploma ARMIET, Maharashtra, India
[3]Student, Dept. of AIML Diploma, ARMIET, Maharashtra, India
[4]Student, Dept. of AIML Diploma, ARMIET, Maharashtra, India
[5]Student, Dept. of AIML Diploma, ARMIET, Maharashtra, India
[6]Student, Dept. of AIML Diploma, ARMIET, Maharashtra, India

---

Abstract -*The power of successful communication cuts across boundaries, cultures, and languages in our increasingly globalized society. However, linguistic boundaries can provide significant challenges that obstruct the free exchange of concepts and understanding. This research sets out to investigate and evaluate a real-time text-to-speech converter and translator system as a pot Text essential means of bridging these linguistic gaps. This in-depth study explores the creation and application of a real-time translation and text-to-speech system, overcoming linguistic obstacles and facilitating efficient communication. It goes through the problem statement, goal, scope, literature review, system architecture, modules, requirements, design, project implementation, analysis of the results, and prospects for the system going forward. The main goal of the project is to develop an adaptable and user-friendly system that enables people to interact easily across language barriers, successfully dismantling the obstacles that frequently impede communication in our multicultural and globally interconnected society. This system prioritizes inclusion and serves a wide range of users, such as visitors, people with language barriers, and companies who conduct international business*.

Key words: *Real Time Text to Speech, Language Barriers, Translation, Effective Communication, Multiple Language Translator*

1. INTRODUCTION

1.1 Problem Statement:

Communication skills are essential in today's fast-paced global world, as they cut beyond geographic and linguistic barriers. However, the difficulties posed by linguistic barriers have grown more acute as the world grows more integrated. As a basic human right, the ability to communicate should not be impeded by language barriers. We describe a solution to this problem that offers real-time text-to-speech conversion and translation using cutting-edge technology. Regardless of the languages they speak, we want to make sure that people can communicate effectively and efficiently with one another.

Linguistic Diversity: There are more than 7,000 languages spoken throughout the world, the linguistic landscape is extraordinarily varied. Numerous languages, dialects, and regional variations are common within a same geographic area. This linguistic variation creates a big problem since it frequently results in misunderstandings, exclusions, and poor communication..

Language Barriers in the Digital Age: Despite the interconnectedness of the digital age, linguistic boundaries still exist in a variety of settings.ಠ Linguistic barriers impede efficient communication for tourists, immigrants, and multinational corporations. Non-native speakers or persons with language barriers may have reduced access to opportunities, services, and information.

The Problem to Solve: In a worldwide culture, addressing the complex issue of linguistic variety and language barriers is the main concern. The objective is to provide a technologically advanced solution that enables people to effortlessly connect with each other, regardless of the languages they speak. Language barriers should no longer prevent people from communicating effectively; instead, they should serve as a bridge to inclusivity and understanding..

1.2 Purpose

The main goal of this project is to create a useful and strong tool that makes cross-language communication easier. Our goal is to enable users be they tourists discovering new cultures, people with limited language skills, or companies conducting business internationally to easily overcome language obstacles. The goal goes beyond simple translation; it includes developing a system that translates text into speech that sounds natural in real time, guaranteeing that the message's subtleties and spirit are accurately communicated.

1.3 Scope

This project has a broad scope with the goal of developing an adaptable and user-friendly software system that can handle text-to-speech conversion and real-time translation for numerous languages. We see a system that supports dialects, lesser-spoken languages, and special linguistic requirements in addition to common languages. It is our goal to make the system adaptable and customizable, so it can cater to the unique communication requirements of diverse user groups. This project endeavors to make communication accessible, intuitive, and inclusive for all..

2. LITERATURE SURVEY

Sr	Author	Year	Title	Working
1	G. K. K. Sanjivani S. Bhabad	2013	Spectral voice conversion for text-to speech synthesis	The OCR and TTS synthesize were actualized to extricate the content data from images and convert it into the sound
2	Kaveri Kamble, ,Ramesh Kagalkar	2012	A Review: Translation Text to Speech Conversation for Hindi Language	In this work text into grammars and then that Grammar into speech is converted by MATLAB. The definition considers are, it does not read punctuation , Romans number
3	P,K. Kurzekar,	2014	A Comparative Study of Feature Extraction Techniques for Speech Recognition System	This approach can efficiently distinguish the object of attention from the background or other objects in the camera view.OCR is used to make word identification on the surrounded text fields and transform it into speech output for blind users.
4	N.K.P.K. Bhupinder Singh	2012	Speech Recognition with Hidden Markov Model: A Review	It includes the correctness of grammar and meaning with end results of achieving excellence in pronunciation
5	Asha G et.al	2017	Image to Speech Conversion for Visually Impaired	The text region from the complex background and to give a high-quality input to the OCR. The text, which is the outcomes of the OCR is given to the TTS engine which provides the speech output.
6	Deepa V.Jose et.al	2014		It includes the correctness of grammar and meaning with end results of achieving

				excellence in pronunciation.
7	Pawan S Nadig	2020	Kannada Text To Speech Conversion System	Optical Character Recognition technique Kannada text images are completely converted into their corresponding speech. The resulting accuracy will be dependent on the training data. Also, use input language as English (India) to get a proper accent.
8	D. Balaji	2018	Implementation of Text To Speech with Natural Voices for Blind People	In this work text into grammars and then that Grammar into speech is converted by MATLAB. The definition considers are, it does not read punctuation, Romans number.
9	K. Lakshmi, T. Chandra Sekhar Rao	2016	Designs and Implementation of Text to Speech Conversion using Raspberry Pi	The translation tools can convert the text to the coveted language and then Again by using the Google speech recognition tool can convert that changed text into voice
10	Ashwini V et.al	2017	An Overview of Technical Progress in Speech Recognition,	A novel text localization algorithm is proposed to localize text regions. Off the-shelf OCR is used to perform text recognition on text localized regions and then recognized text codes are transforme d to speech for a blind person.

3. PROPOSED SYSTEM

3.1 System Architecture: The proposed system's architecture is designed to be robust and efficient. It follows a client server model where the client sends text input, and the server processes this input, converting it to speech and translating it as necessary. The architecture includes several key components, such as:

tkinter : This is the standard GUI (Graphical User Interface) toolkit for Python. It provides various widgets and methods to create simple and complex GUI applications.

tkt from tkinter : ttk is a submodule of tkinter, providing themed widget set. It includes classes that provide a more modern look and additional functionalities.

gtts : gtts stands for "Google Text-to-Speech". It is a Python library and CLI tool to interface with Google Translate's text-to-speech API. It allows the conversion of text to speech.

subprocess (import subprocess): This module allows you to spawn new processes, connect to their input/output and

obtain their return codes. In this code, it's used to run the system's default audio player to play the generated audio file.

googletrans: googletrans is the Python wrapper for Google Translate API. It allows you to translate the text from one language to another language.

tkinter messagebox: This module provides a set of standard dialog boxes that pop up over your. Tkinter application's main window. It's used here to show error messages in case of exceptions.

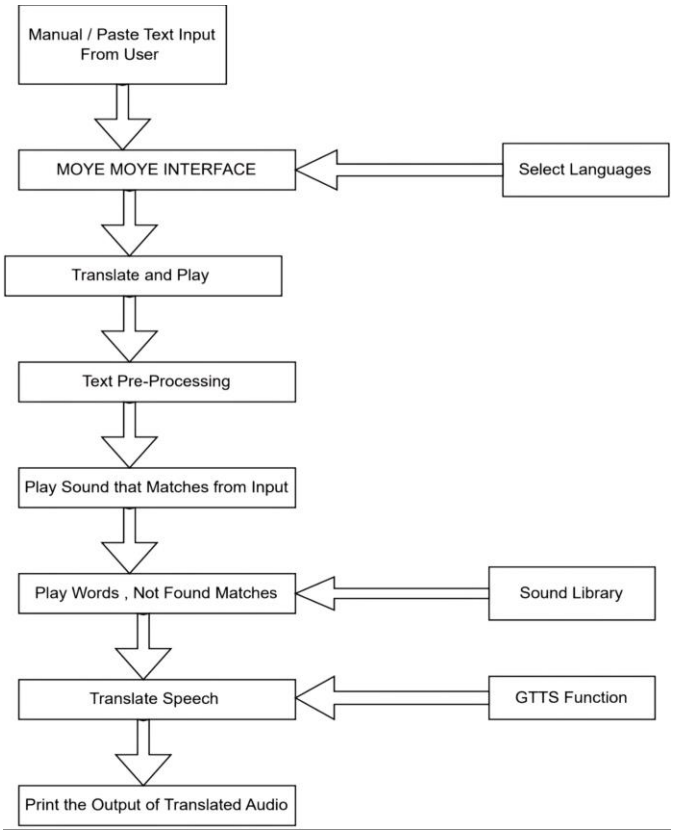


FIG 3.1 SYSTEM ARCHETECTURE

4. SYSTEM REQUIREMENTS

4.1 Software Requirements

Operating System: The code is compatible with any operating system that supports Python and the required libraries.

Python Interpreter: Download and install Python from the official website: python.org

Python Libraries: tkinter, gtts (Google Text-to-Speech) Google trans, tkinter messagebox ,subprocess

Internet Connection: Required for accessing Google Translate API through the googletrans library.

Audio Output Device: A functioning audio output device such as speakers or headphones is required to listen to the translated text.

4.2 Hardware Requirements

Computer: Any modern computer or laptop is sufficient.

Processor: Intel Core i3 or AMD RYZEN 5 equivalent (or higher) recommended.

RAM: Minimum 4GB (8GB or higher recommended) for smooth performance.

Operating System: The code can run on any operating system supported by Python and its libraries (e.g., Windows, macOS, Linux).

Internet Connection: An active internet connection is required for accessing the Google Translate API through the googletrans library.

Audio Output Device: A functioning audio output device such as speakers or headphones is necessary to listen to the translated text.

Server: The server component requires a high-performance CPU with multiple cores, ample RAM for handling concurrent requests, and significant storage capacity for storing language models, translation data, and voice databases.

Client: The client-side can be accessed on standard personal computers or mobile devices with internet connectivity. There are no stringent hardware requirements for users, as most modern devices can efficiently run the user interface

4.3 Technology Used

Python: Python is the primary programming language used for developing the application**.**

Tkinter: Tkinter is a standard GUI (Graphical User Interface) toolkit for Python.It provides various widgets and methods to create simple and complex GUI applications. Tkinter is used for building the graphical user interface of the application.

gtts (Google Text-to-Speech): gtts is a Python library and CLI tool to interface with Google Translates text-to-speech API.It allows the conversion of text to speech. Used for generating speech from translated text.

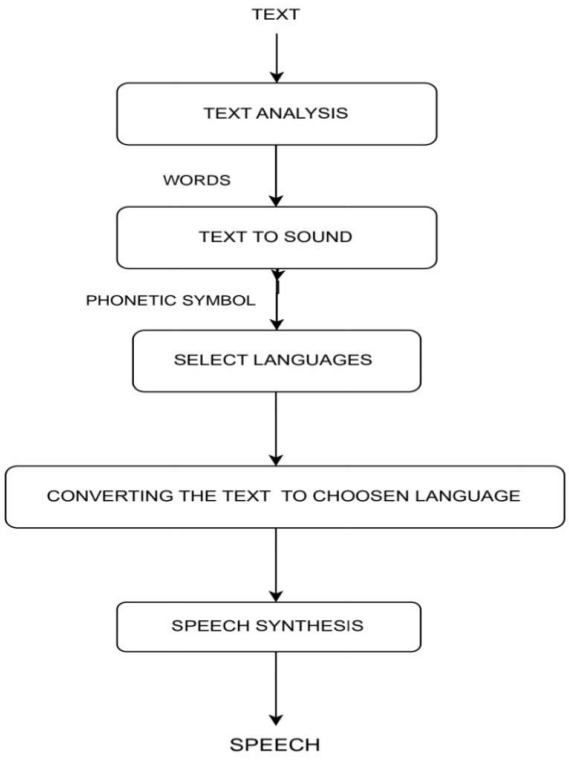
googletrans: googletrans is a Python wrapper for Google Translate API.It provides functionalities to translate text from one language to another. Used for real-time translation of user input. sand machine learning models trained on large corpora of text in different languages.

subprocess: The subprocess module allows spawning new processes, connecting to their input/output/error pipes, and obtaining their return codes. Used for running the system's default audio player to play the generated audio file**.**

Natural Language Processing (NLP): NLP is at the core of language detection, translation, and speech synthesis. Machine learning models, including recurrent neural networks (RNNs) and transformers, play a vital role in language analysis and translation.

5. SYSTEM DESIGN

5.1 UML Diagram:



UML DIAGRAM 5.1

To illustrate the system's architecture and the interactions between different modules, Unified Modeling Language (UML) diagrams are employed.

6.PROJECT IMPLEMENTATION

6.1 Algorithm Used for Implementation:

Language Detection Algorithm: The language detection in the provided code is facilitated by the Google Translate API through the google trans library. While the exact algorithm used by Google for language detection is proprietary, it likely involves statistical method sand machine learning models trained on large corpora of text in different languages.

Text-to-Speech Algorithm: The text-to-speech conversion is handled by the Google Text-to-Speech (gtts) library. The specific algorithm used by Google for text-to-speech synthesis is not disclosed, but it likely involves neural network-based models trained on large datasets of human speech**.** The gtts library sends the text to Google's servers, where the text is synthesized into speech using Google's advanced algorithms.

Translation Algorithm: The translation module in the provided code also utilizes the Google Translate API through the googletrans library. The Google Translate API employs sophisticated machine learning models, including neural machine translation (NMT) models, for translation. These models consider the context and semantics of the text to provide accurate translations between languages. While the exact architecture of the models used by Google is proprietary, it is known that they leverage transformer architectures and are trained on vast amounts of bilingual text data.

7.CONCLUSION

In this project we have made an effort to understand how, in the long run, a changing world can dissolve linguistic barriers. We have spent a lot of time investigating this fascinating real-time text-to-speech and translation system, which looks like it could revolutionize how difficult it is to overcome language hurdles. We've delved into the nature of the issue, established some objectives, and determined what equipment we require through our investigation. From a practical standpoint, we have considered how this entire situation might develop in the future. It's similar to moving toward a future in which everyone may click to connect, regardless of their language or place of origin, increasing the sense of awesomeness and community throughout the world. In a world where connectivity knows no bounds, effective communication has become the goal of our global society. The ability to transcend language barriers and foster understanding among diverse cultures is an imperative endeavor. The real-time text-to-speech converter and translator system presented in this report represents that not only a technological achievement but a testament to our unwavering commitment to breaking down linguistic divides. The main purpose of this project is to develop a system that empowers individuals to communicate seamlessly across languages. It is a system built on the principles of accessibility, inclusivity, and user-friendly interface.

8.REFERENCES

[1] G. K. K. Sanjivani S. Bhabad, "An Overview of Technical Progress in Speech Recognition," International Journal of Advanced Research in Computer Science and Software Engineering, 2013.

[2] Kaveri Kamble, Ramesh Kagalkar,"A Review: Translation Text to Speech Conversation for Hindi Language" (IJSR) 2012

[3] P, K. Kurzekar, "A Comparative Study of Feature Extraction Techniques for Specch Recognition System," International Journal of Innovative Research in Science. Engineering and Technology. 2013

[4] N. K. P.K. Bhupinder Singh, ""Speech Recognition with Hidden Markov Model: A Review," International Journal of Advanced Research in Computer Science and Software Engineering, 2012

[5] Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik (2017) "Image to Speech Conversion for Visually Impaired", International Journal of Latest Research in Engineering and Technology, ISSN: 2454 5031, Issue 06, Vol. 03, No. 0, pp. 09-15.

[6] Deepa V.Jose, Alfateh Mustafa, Sharan R(2014) "A Novel Model for Speech To Text Conversion", International Refereed Journal of Engineering and Science, ISSN(online): 2319-183X,(print) 2319-1821, Issue 1, Vol. 3, No. 0, pp. 39-41.

[7] Pawan S Nadig, Pooja G, Kavya D, R Chaithra, Radhika A D (2020) "Kannada Text To Speech Conversion System", International Journal of Innovative Technology and Exploring Engineering, ISSN: 2278-3075, Issue 10, Vol. 9, No. 0, pp. 0.

[8] D. Balaji, M. Rajukumar, P. Sethu, V. Hariharan, S. Sekar (2018) "Implementation of Text To Speech with Natural Voices for Blind People", Paideuma, ISSN: 0090 5674, Issue 2, Vol. 11, No. 0, pp. 49-53.

[9] K. Lakshmi, T. Chandra Sekhar Rao (2016) "Designs and Implementation of Text to Speech Conversion using Raspberry Pi", International Journal of Innovative Technology and Research, Issue 6, Vol. 4, No. 0, pp. 4564 4567.

[10] Ashwini V, Mhaske, Mahesh S. Sadavarte, R Chaithra, Radhika A D (2016) "Portable Camera Based Assistive Text Reading from Hand Held Objects for Blind Person", International Journal of Advanced Research in Computer and Communication Engineering, ISSN (print): 2319 5940, ISSN(online): 2278-1021, Issue 6, Vol. 5, No. 0, pp. 0.