A
Mini-Project Report on

AMPYtranslate

Submitted in complete fulfillment of the requirements
for the degree of
BACHELOR OF ENGINEERING
IN
Computer Science & Engineering
Artificial Intelligence & Machine Learning

By

Manasi Patil (22106042)
Aniruddha Pawar (22106009)
Yash Penkar (22106118)
Prarthana Patil (22106035)

Under the guidance of

Prof. Ranjita Asati

**Department of Computer Science & Engineering
(Artificial Intelligence & Machine Learning)
A. P. Shah Institute of Technology
G. B. Road, Kasarvadavali, Thane (W) - 400615
University Of Mumbai
2023-2024**

A. P. SHAH INSTITUTE OF TECHNOLOGY

CERTIFICATE

This is to certify that the project entitled "**AMPYtranslate**" is a bonafide work of Manasi Patil (22106042), Aniruddha Pawar (22106009), Yash Penkar (22106118), Prarthana Patil (22106035) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of **Bachelor of Engineering** in **Computer Science & Engineering (Artificial Intelligence & Machine Learning).**

Prof. Ranjita Asati	Dr. Jaya Gupta
Mini Project Guide	Head of Department

A. P. SHAH INSTITUTE OF TECHNOLOGY

Project Report Approval

This Mini project report entitled "**AMPYtranslate**" by **Manasi Patil, Aniruddha Pawar, Yash Penkar and Prarthana Patil** is approved for the degree of *Bachelor of Engineering* in *Computer Science &Engineering*, (AI&ML) *2023-24*.

External Examiner: _____

Internal Examiner: _____

Place: APSIT, Thane
Date:

Declaration

We declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission hasnot been taken when needed.

Manasi Patil   Aniruddha Pawar   Yash Penkar   Prarthana Patil

(22106042)   (22106009)   (22106118)   (22106035)
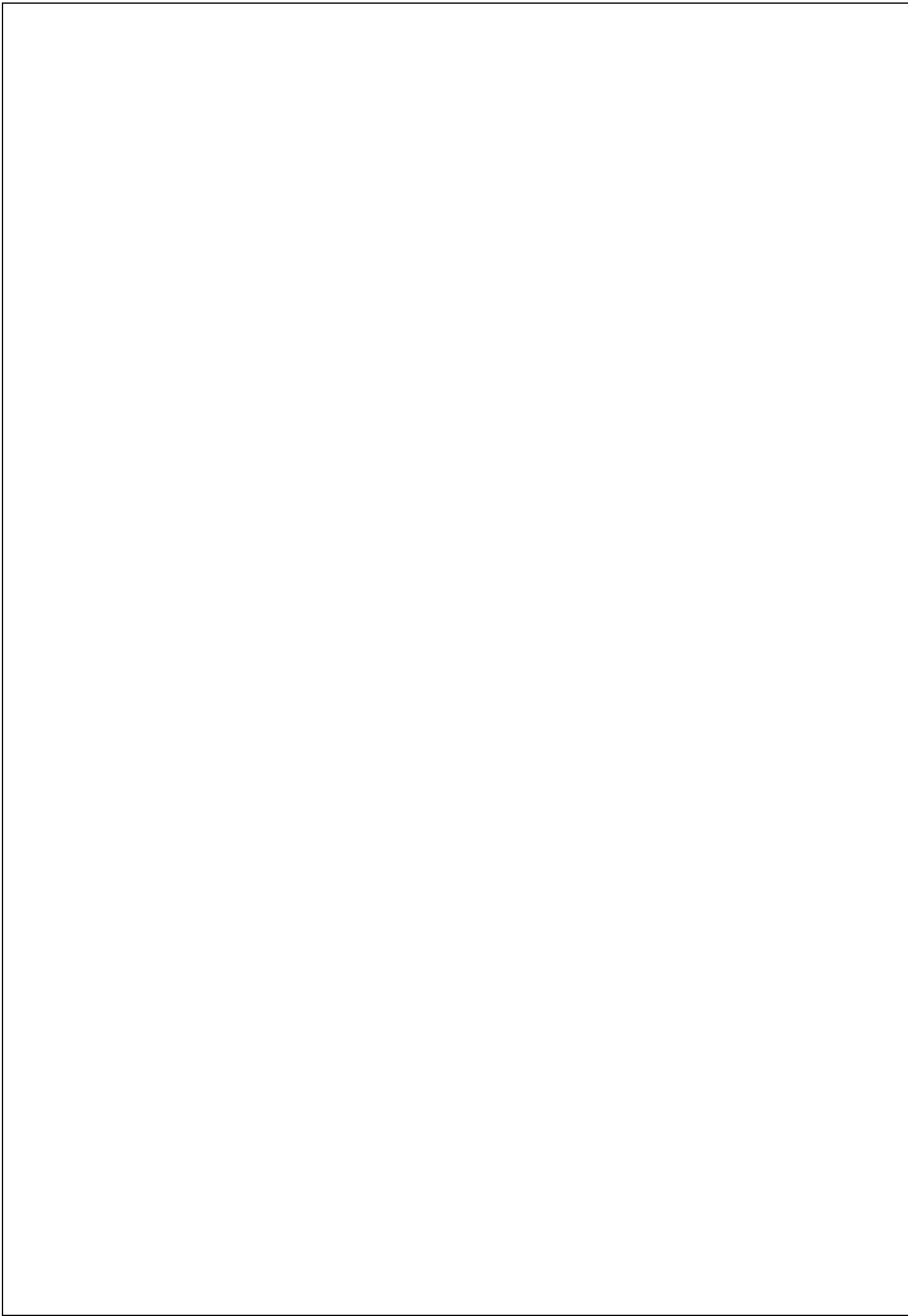
ABSTRACT

A language translator is a web that can be utilized for translating from one language to another. The problem of language difference has hindered effective information communication over the years. This traditional approach used for solving the problem of language differences has not been productive and favorable. The study develops an website in order to make learning and language translation easy and facilitates stress-free communication. The system will also be able to evaluate language translation to determine their suitability for everyday conversation. Website to translate from English language to French, German, Hindi, etc and vice versa. In addition, many other language translations have been added namely Marathi to Mandarin and Tamil. It provides categories of commonly used words and phrases. As such, tourists can rely on mobile phone for communication purposes. Website dictionary allows for word and phrases search as a better traveling guide. This will have the functionality of pronouncing the words and phrases.

INDEX

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

A voice-based language translator is a technological solution that enables real-time conversion of spoken language from one language to another. It leverages advanced natural language processing and machine learning techniques to recognize and interpret spoken words and phrases, translating them into the desired target language. This technology allows individuals to communicate with others who speak different languages without the need for manual translation or interpretation.

Voice-based language translators often come in the form of mobile apps, devices, or software platforms or websites. They have become increasingly popular in a globalized world, facilitating cross-cultural communication, travel, and business interactions. These systems use voice recognition to capture the source language and employ sophisticated algorithms to generate accurate translations in near-real-time, making it easier for people from diverse linguistic backgrounds to understand and engage in meaningful conversations. A voice-based language translator is a sophisticated tool that harnesses the power of speech recognition and machine translation technologies to enable seamless communication between individuals who speak different languages. These systems start by capturing the spoken words or phrases of the user. This involves converting audio input (your voice) into text that the software can work with. Advanced speech recognition algorithms analyse the acoustic features of the speech and convert them into written words. Once the spoken words are transcribed into text, the system identifies the source language. This step is crucial because it determines which language the software needs to translate from. After identifying the source language, the system employs machine translation algorithms to convert the text from the source language into the desired target language. Machine translation models, often powered by neural networks, have significantly improved the accuracy of translations in recent years.

These systems also rely on natural language processing (NLP) techniques to understand context, idiomatic expressions, and grammar. NLP helps ensure that translations are not only accurate but also sound natural to native speakers.

Once the translation is generated, the system converts the translated text back into spoken words in the target language. This allows the user to hear the translation through the device or app, making it a truly voice-based experience.

Voice-based language translators are designed for real-time communication. Users can speak

sentences or phrases, and the device or app will quickly provide translations, enabling back-and-forth conversations in different languages without significant delays.

To make these tools accessible to a wide range of users, they often come with user-friendly interfaces, such as mobile apps or dedicated devices with touchscreens or buttons. Some even offer visual translations on screen, which can be especially useful in noisy environments.

Voice-based language translators require an internet connection to access large language databases and perform accurate translations.

Overall, voice-based language translators have revolutionized cross-cultural communication, making it more convenient and accessible for people to connect, travel, and do business with individuals who speak different languages. They continue to evolve with advancements in AI and NLP, offering increasingly accurate and natural translations.

CHAPTER 2
LITERATURE SURVEY

2. LITERATURE SURVEY

2.1-HISTORY

The evolution of voice-based language translation is a remarkable journey through the development of technology, with roots dating back to the mid-20th century. Early research in machine translation during the 1950s and 1960s laid the foundation, though these efforts yielded rudimentary and often inaccurate results. Simultaneously, the field of speech recognition was taking shape, with the emergence of systems capable of converting spoken words into text. These early speech recognition systems, however, were far from perfect. As the years passed, rule-based translation systems emerged in the 1960s and 1970s, utilizing linguistic rules and dictionaries for translations. Although they represented progress, these systems struggled with the complexities of natural language. In the 1990s, a shift towards statistical machine translation (SMT) occurred, leveraging statistical algorithms to improve translation accuracy. The convergence of speech recognition and machine translation technologies in the early 2000s paved the way for voice-based language translation systems. These systems enabled users to speak sentences in one language and receive real-time spoken translations, marking a significant milestone. The proliferation of smartphones and high-speed internet in the 2010s led to the development of mobile apps offering voice translation capabilities, making this technology accessible to a wider audience. Notably, neural machine translation (NMT) models, driven by artificial intelligence and deep learning, have revolutionized translation accuracy in recent years. These models, integrating context-aware and fluent translations, have been incorporated into voice-based translators, significantly enhancing their performance. Additionally, smart speakers and home assistants like Amazon Echo and Google Home have integrated voice-based translation capabilities. Today, voice-based language translation continues to evolve, incorporating state-of-the-art natural language processing and machine learning techniques. It supports an expanding array of languages, provides real-time, natural-sounding translations, and holds promise for further advancements, ensuring that language barriers become increasingly surmountable in our interconnected world.

2.2-LITERATURE REVIEW

[1] Words Echoing Through Time: The Melody of Learning with A Voice Translator App Article: *Article in Pulchra Lingua a Journal of Language Study Literature & Linguistics · November 2022.*

The study investigated the effectiveness of using a voice translator application to enhance students' vocabulary acquisition and interest in learning. The researcher collected both qualitative and quantitative data through a questionnaire and vocabulary test to analyze students' experiences and perceptions of using the application. The findings showed that the voice translator application had a positive impact on students' vocabulary acquisition and interest in learning. The qualitative data indicated that students found the application helpful for learning vocabulary, while the quantitative data showed a significant improvement in their vocabulary scores.

[2] Speech Recorder and Translator using Google Cloud Speech-to-Text and Translation: *J.Y. Chan and H.H. Wang 1, Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Journal of IT in Asia, Vol 09, 2021*

The web-based speech recorder and translator is successfully developed. It aims to increase the audience's accessibility to the video content that is using a language that the audiences do not know. The application is functioning well except it is limited to the quota of the Google Cloud STT and Translation plan. In conclusion, the application is successfully developed with the functions of detecting audio, generating text, and translating text between English and Mandarin.

[3] Analysis of Voice Recognition System on Translator for Daily Use: *Yudi Aryatama Fonggi1*, Tio Oktavianus JURNAL EMACS (Engineering, MAthematics and Computer Science) Vol.3 No.2 May 2021.*

The use of voice recognition system translator devices is commuted, has a positive impact, and can still be developed in various aspects, the use of voice recognition in multilingual communication really helps the translation process between different languages, especially communication during the same period both online and offline using a portable device. The basic algorithm used in the voice recognition translator is to use a voice recorder as a receiving device for the user's voice input then it will be matched with the language that has been

inputted in the system then the voice signal will be translated according to the desired language in the database system or by a second user as the first user partner.

[4] Investigative Study towards the Development of Mobile Language Translator Art:
International Journal of Digital Content Technology and its Applications · December 2012.
The research's primary objective is to create a mobile application designed to address communication barriers among tourists. In conclusion, the development of this mobile-based system has proven to be a time-intensive endeavor, aiming to deliver a comprehensive application with enhanced features and functionalities. However, certain limitations stemming from the application programming interface (API) have impacted the development process. Based on the research findings, several recommendations can be made to better cater to the needs and requirements of end-users.

[5] Language Translator Application: *M Vaishnavi1, HR Dhanush Datta2, Varsha Vemuri3, L Jahnavi4 1, 2, 3, 4B.E Student, Dept of CSE, Ballari Institute Of Technology and Management, Ballari, Karnataka, India. Volume 10 Issue VII July 2022.*
In the existing system, we have google translator which utilizes internet connectivity whereas internet may not be available all the time and there are also many android applications available that may not support all the functionalities like scanning text, speech recognition and translates the text and which are applicable for specific and limited languages which are not useful for all the users. So here in the proposed system where we will be implementing translation with support all the functionalities like scanning text, speech recognition and translates the text and includes the languages which are popular in our country as well as popular all over the world. The advantage of this application is it doesn't require internet connectivity.

[6] Voice to Voice Language Translation System: *International Journal of Engineering Research & Technology (IJERT) IJERT ISSN: 2278-0181 IJERTV3IS100924 www.ijert.org (This work is licensed under a Creative Commons Attribution 4.0 International License.) Vol. 3 Issue 10, October- 2014*
Voice to Voice Language Translation system is a device that is designed to bridge the language gap between individuals and foreigners when traveling in our country. The need arises from the inability of dictionaries and human translators to suit our needs for better communication. At present we need "Personalized Interpreters" which will reduce our dependence on

dictionaries and human interpreters. This will reduce the hindrance posed by the language barrier. In this situation the system proposed will suffice the purpose reasonably well and minimize the communication inefficiencies. The system can overcome the real time difficulties of illiterate people and improve their lifestyle.

[7] Study on the Use of Speech Recognition Function to Practice Speaking English Using the Voice Translator "Pocketalk": *Harumi Kashiwagi, Kobe University, Japan Min Kang, Kobe University, Japan Kazuhiro Ohtsuki, Kobe University, Japan. The Asian Conference on Education & International Development 2021 Official Conference Proceedings.*

The study explored the use of speech recognition to support the practice of English speaking by using the voice translator Pocketalk. They conducted a study in which English sentences spoken by 95 Japanese university students were identified by the speech recognition function of this device. Afterward, a five-point Likert scale was used to measure the usefulness of the activity with Pocketalk and the affective factors related to speaking English.

[8] Voice interfaces for real-time translation of common tourist conversation: *Conference Paper · January 2011*

In this paper, they presented the evolution of a voice interface real-time translation interaction style. The design requirements were identified in the literature and after applying individual qualitative interviews. They could evolve the concept model in three design interactions involving users. Also evaluated each version considering the user's feedback to perform the task analysis. As a result, a voice interface for real-time translator was developed.

[9] Direct Speech to Speech Translation Using Machine Learning: *Sireesh Haang Limbu December 2020.*

The results of the experiments conducted on the three models explained in the above section have been broadly divided into two categories: The first one using the high-quality human audio files taken from the '50 languages' website explained in the 'Data Representation' section. The second subsection describes the results procured using the 'Opus' parallel text corpus that have been converted to audio using Google TextTo-Speech API.

CHAPTER 3

Problem Statement

3. PROBLEM STATEMENT

One of the most obvious problems was the difficulty in communicating with individuals who spoke different languages. Language Barriers was especially problematic for travelers, tourists, and individuals doing business in foreign countries. Attempting to communicate in a language one did not fully understand often led to misunderstandings, which could have consequences ranging from minor inconveniences to serious issues in professional or personal interactions.

The inability to communicate effectively limited social interactions, cultural exchange, and the sharing of knowledge across linguistic boundaries.

For international business, diplomacy, and academia, reliance on human interpreters and translators could be time-consuming and costly. Learning a new language was often the only solution to overcome language barriers. This required significant time and effort, making it less accessible to many.

CHAPTER 4

Experimental Setup

4.EXPERIMENTAL SETUP

4.1 HARDWARE SETUP

The system configuration required to run the software includes the following:

1. Operating System: Compatible with Windows, macOS, and Linux.

2. Hardware: Standard CPU, 4GB+ RAM, sufficient storage.

3. Python 3: Installed and added to PATH.

4. Internet Connection: Required for online services.

5. Audio: Functional microphone and audio output.

6. Permissions: To access microphone and file system.

7. Dependencies: Install required Python packages.

8. Sound Configuration: Verify system audio settings.

9. Language Support: Ensure system supports languages.

10. Firewall/Network: Ensure access to external resources.

11. Driver Compatibility: Up-to-date microphone and audio drivers.

12. CPU Performance: Better CPUs for intensive use.

These requirements should help set up and run the software effectively on the system.

4.2 SOFTWARE SETUP

1. **Operating System**: Software often requires a specific operating system (e.g., Windows, macOS, Linux) and a minimum version of that operating system.

2. **Processor**: The software may have minimum processor requirements, such as a specific type (e.g., Intel Core i3)

3. **Memory (RAM)**: The software may require a certain amount of RAM for optimal performance. For example, it might specify a minimum of 4 GB .

4**. Storage Space**: Software typically needs a certain amount of available storage space on the computer's hard drive or SSD.

5**. Display Resolution**: Certain software may have minimum display resolution requirements to ensure proper functionality and optimal user experience.

6. **Additional Software Dependencies**: Some software may require specific additional software or libraries to be installed beforehand. These dependencies might include runtime environments, drivers, or other software components.

7**. Internet Connection**: Certain software may require an internet connection for installation or to access online features and updates.

CHAPTER 5

Proposed System & Implementation

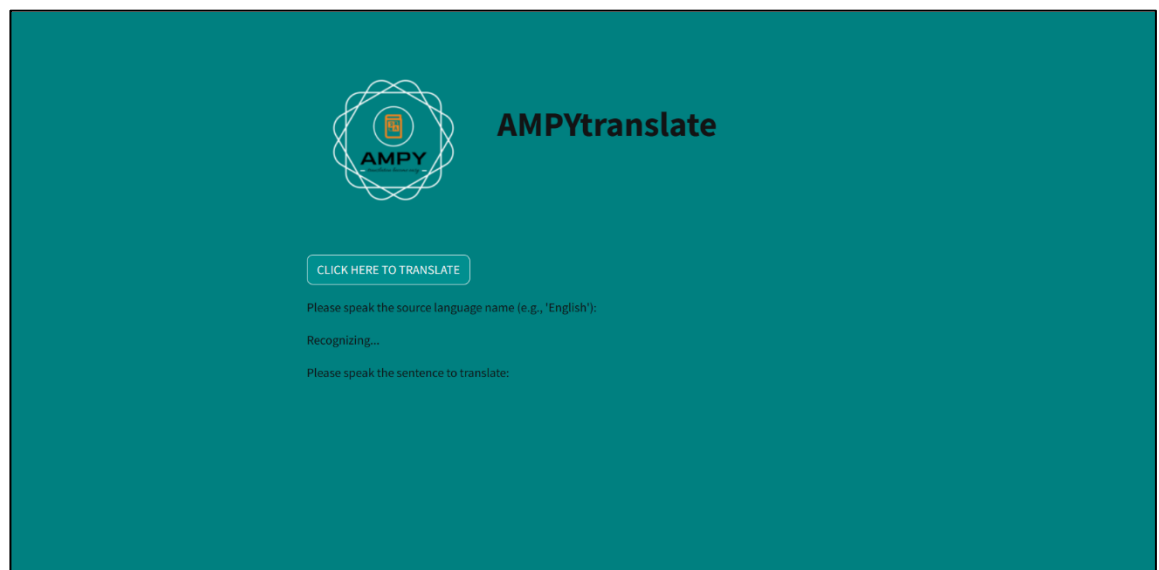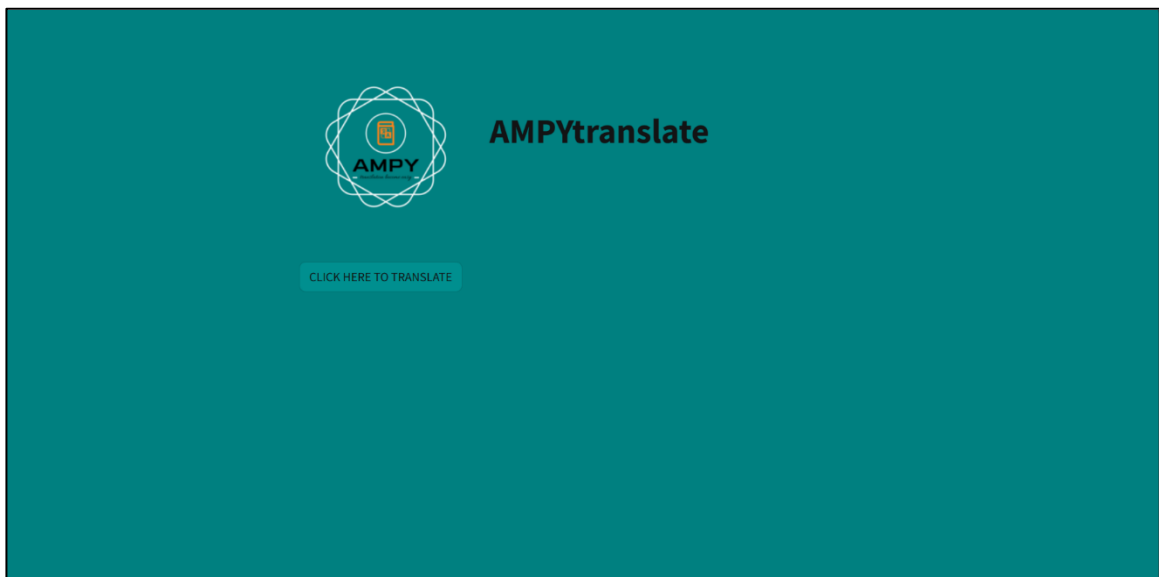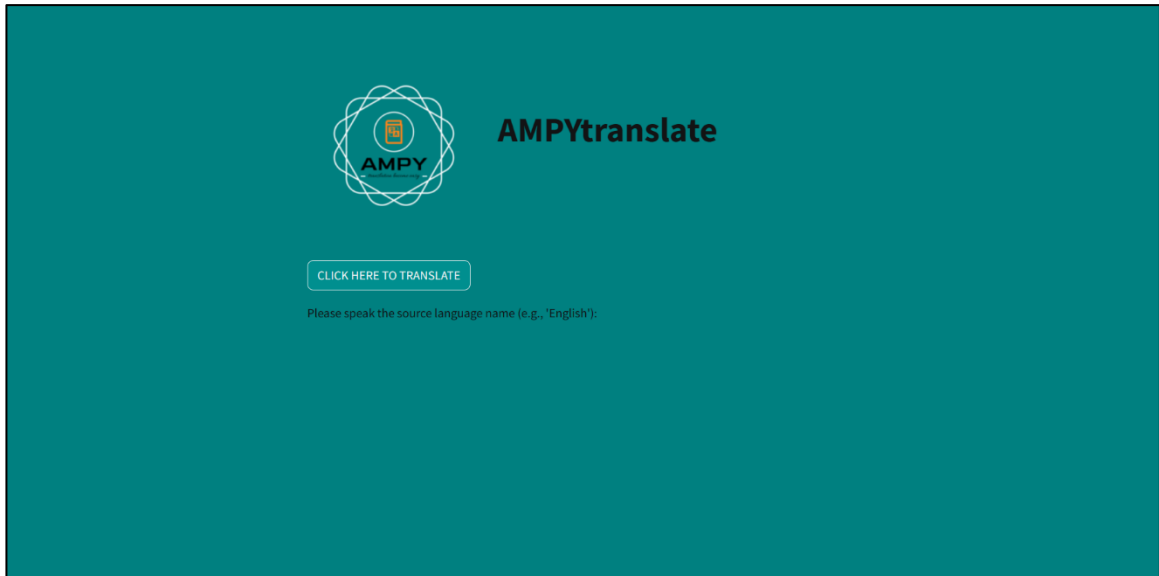# 1. PROPOSED SYSTEM & IMPLEMENTATION

5.1 BLOCK DIAGRAM OF PROPOSED SYSTEM
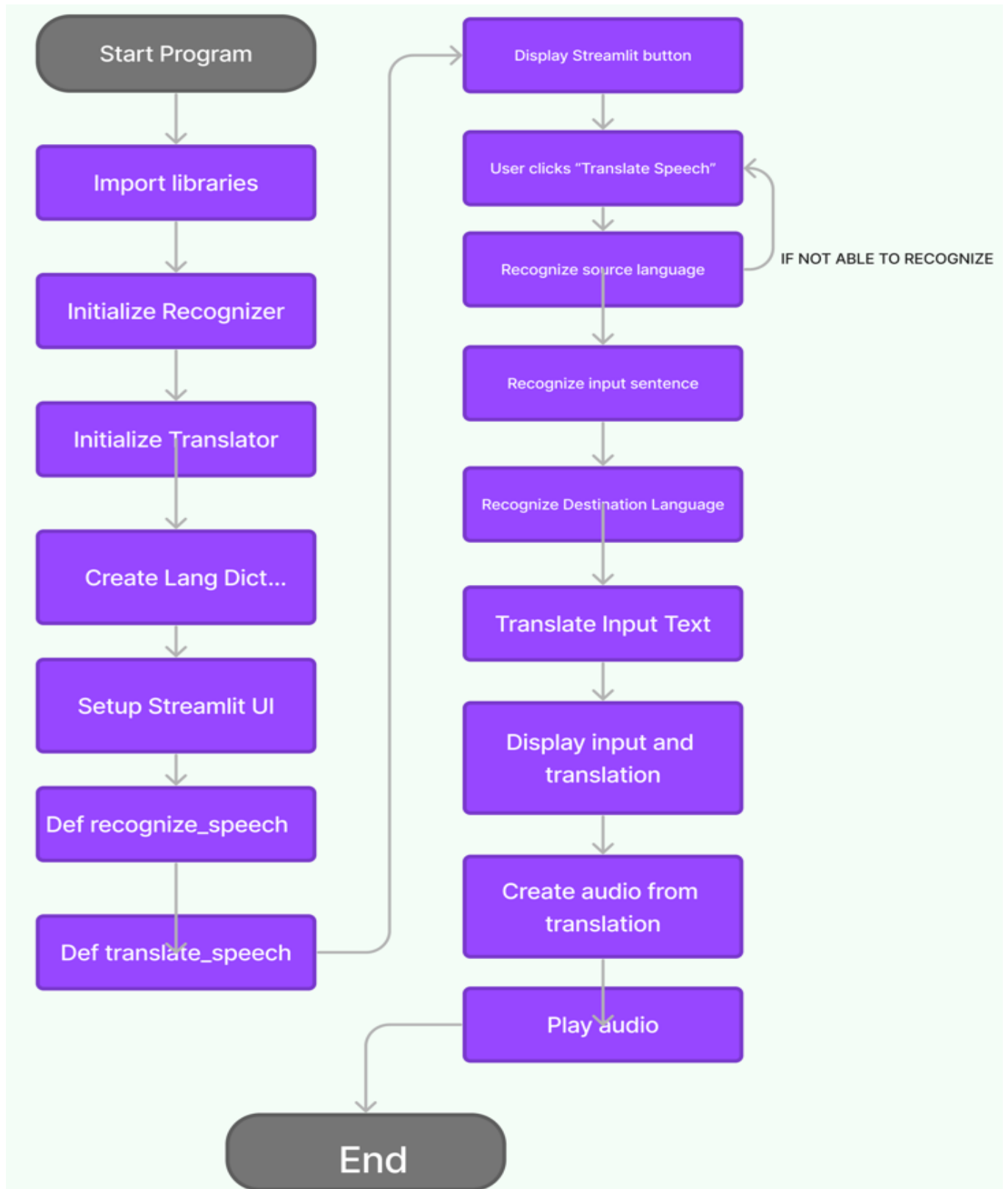
5.2 DESCRIPTION OF BLOCK DIAGRAM

1. **Import Program**: The program begins.
2. **Start libraries**: The necessary libraries for the program are started.
3. **Initialize Recognizer**: The recognizer, which might be used for speech recognition, is initialized.
4. **Initialize Translator**: The translator, possibly used for language translation, is initialized.
5. **Create Lang Dict**: A language dictionary is created, likely to support multiple languages.
6. **Setup Recognize_UI**: The user interface for the recognizer is set up.
7. **Def translate_speech**: A function named 'translate_speech' is defined.
8. **User clicks "Translate" button**: The user interacts with the program by clicking the "Translate" button.
9. **Recognize language IF NOT ABLE**: If the program is unable to recognize the language, this step is executed.
10. **Recognize input language**: The program recognizes the language of the input provided by the user.
11. **Translate input text**: The recognized text is then translated.
12. **Display input**: Finally, the translated input is displayed.

5.3 IMPLEMENTATION

5.4 ADVANTAGES

The software offers a wide range of advantages, making it a versatile and valuable tool:

1. **Speech Recognition:** The software enables easy input through speech recognition, making it user-friendly and accessible.

2. **Real-Time Language Translation**: It provides real-time language translation, making it a valuable tool for multilingual communication.

3. **User-Friendly Interface**: The Streamlit interface ensures a straightforward and intuitive user experience.

4. **Customization**: Users can customize the software for various source and destination languages to meet their specific needs.

5. **Text-to-Speech Capability**: The software includes text-to-speech functionality, allowing users to hear the translated text, enhancing the communication experience.

6. **Audio Playback**: Users can play back the translated text as audio, improving the accessibility and user experience further.

7. **Convenience for Multitasking**: The software allows users to convert spoken language to text and vice versa, which is particularly useful when multitasking or when typing isn't practical.

8. **Language Learning Support**: It can aid language learners by providing both text and audio in their target language, promoting better language comprehension and pronunciation.

9. **Cost-Effective Communication**: It's a cost-effective solution for real-time language communication, reducing the need for expensive human translators or language learning courses.

10. **Global Accessibility**: The software is accessible worldwide, making it an ideal tool for connecting people from different linguistic backgrounds, both professionally and personally.

11. **Fast and Efficient Communication**: Real-time translation and audio playback enable efficient and timely communication, reducing language barriers and misunderstandings.

12. **Versatile Communication Tool**: It can be used in various settings, from business meetings and international conferences to everyday conversations and travel situations.

CHAPTER 6

Conclusion

6. CONCLUSION

The "AMPYtranslate" software represents a significant step forward in the realm of language technology and communication. By bridging language barriers and providing seamless multilingual communication, this software has the potential to revolutionize the way we interact and connect with one another. The software's list of advantages is impressive and encompasses various features, from speech recognition to real-time language translation and text-to-speech capabilities. It is a powerful tool that offers practical solutions for a multitude of scenarios. Its user-friendly interface ensures that users of all backgrounds and technical expertise can navigate it with ease. The ability to customize source and destination languages makes it adaptable to individual needs, while the inclusion of audio playback ensures that communication is not only accessible but also engaging and immersive.

6.1 FUTURE SCOPE

The software's capabilities are just the beginning, and there are several potential avenues for further development and improvement:

- Enhanced Language Support: Expand the range of supported languages to make the software even more inclusive and versatile.

- Advanced Speech Recognition: Integrate more advanced speech recognition algorithms and models to improve accuracy and recognition in various accents and dialects.

- Machine Learning Integration: Implement machine learning techniques to enhance the software's translation accuracy and fluency.

- Voice Commands: Enable voice commands for controlling the software, making it hands-free and even more convenient.

- Mobile Application: Develop a mobile application version for on-the-go language translation and communication.

- Cross-Platform Compatibility: Ensure compatibility with different operating systems and platforms to reach a wider user base.

- Customization Features: Allow users to personalize the user interface and settings to meet their specific needs.

REFERENCES

[1] "Words Echoing Through Time: The Melody of Learning with A Voice Translator App" Article in Pulchra Lingua a Journal of Language Study Literature & Linguistics · November 2022

[2] "Speech Recorder and Translator using Google Cloud Speech-to-Text and Translation", 1 J.Y. Chan and 2 H.H. Wang 1, 2 Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia email: 1 jietying@outlook.com, 2 hhwang@unimas.my

[3] "Analysis of Voice Recognition System on Translator for Daily Use", Yudi Aryatama Fonggi1*, Tio Oktavianus2 1,2 Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480

[4] "Investigative Study towards the Development of Mobile Language Translator" Article in International Journal of Digital Content Technology and its Applications · December 2012

[5] "Language Translator Application", M Vaishnavi1 , HR Dhanush Datta2 , Varsha Vemuri3 , L Jahnavi4 1, 2, 3, 4B.E Student, Dept of CSE, Ballari Institute Of Technology and Management, Ballari, Karnataka, India

[6] "Voice to Voice Language Translation System": International Journal of Engineering Research & Technology (IJERT) IJERT ISSN: 2278-0181 IJERTV3IS100924 www.ijert.org (This work is licensed under a Creative Commons Attribution 4.0 International License.) Vol. 3 Issue 10, October- 2014

[7] "Study on the Use of Speech Recognition Function to Practice Speaking English Using the Voice Translator", "Pocketalk" Harumi Kashiwagi, Kobe University, Japan Min Kang, Kobe University, Japan Kazuhiro Ohtsuki, Kobe University, Japan

[8] "Direct Speech to Speech Translation Using Machine Learning", Sireesh Haang Limbu, Uppsala University .