**TITLE: LANGUAGE TRANSLATION VIA SPEECH**

**ABSTRACT:**

This project proposes the development of a Speech-Based Language Translation System that enables real-time translation of spoken language from one language to another. The system integrates Automatic Speech Recognition (ASR), Machine Translation (MT), and Text-to-Speech (TTS) technologies to create a seamless, voice-driven translation experience.

The project leverages advanced tools and APIs such as **OpenAI Whisper**, **Google Translate API**, and **Google Text-to-Speech**, and is developed using a combination of **Python** and **cloud-based services**.

**INTRODUCTION:**

Language is one of the most fundamental tools of human interaction. However, with thousands of languages spoken around the world, communication across linguistic boundaries remains a significant challenge. In an increasingly globalized society, the need for efficient and accessible language translation tools has never been greater. Traditional translation methods, such as dictionaries or text-based translation apps, often fall short when real-time, hands-free, and natural communication is required.

This project focuses on developing a Speech-Based Language Translation System that enables users to communicate across different languages through spoken input. By combining Automatic Speech Recognition (ASR), Machine Translation (MT), and Text-to-Speech (TTS) technologies, the system is designed to take a spoken phrase in one language, translate it accurately into another language, and vocalize the translated output.

**SYSTEM COMPONENTS:**

a. Speech Recognition (ASR)

* Purpose: Converts spoken input into text.
* Tool Used: OpenAI Whisper / Google Speech-to-Text API
* Input: User's voice (e.g., “Hello, how are you?”)
* Output: Transcribed text (“Hello, how are you?”)

b. Machine Translation (MT)

* Purpose: Translates the transcribed text into the target language.
* Tool Used: Google Translate API / DeepL API
* Input: Source language text
* Output: Translated text in the target language (e.g., “Hola, ¿cómo estás?”)

c. Text-to-Speech (TTS)

* Purpose: Converts translated text into natural-sounding speech.
* Tool Used: Google Text-to-Speech / Amazon Polly
* Input: Translated text
* Output: Spoken audio in the target language.

**KEY FEATURES:**

* Real-time speech translation
* Support for multiple languages
* Modular design for easy upgrades
* Can be extended to mobile or IoT devices
* Suitable for various applications: travel, healthcare, education, and more.

**TECHNOLOGIES USED:**

* **Programming Language:** HTML
* **Libraries/Frameworks:**
  + speech recognition / OpenAI-whisper
  + GTTS (Google Text-to-Speech)
  + google trans / official Google Translate API
  + py dub or play sound for audio playback
* **Platforms:** Can be extended to mobile (Android/iOS) or web (React, Flask, etc.)
* **Optional:** Real-time communication using sockets or WebRTC.

LANGUAGE SUPPORT COMPONENTS:

| **Feature** | **Technology Used** | **Role in Language Handling** |
| --- | --- | --- |
| Voice Input | Speech Recognition API (Web Speech API) | Converts spoken language to text. Supports multiple language codes (e.g., ta-IN, ko, ar). |
| Translation | My Memory API (via fetch) | Translates text between languages (based on ISO language codes). |
| Voice Output | Speech Synthesis API | Reads out translated text using supported system/browser voices for each language. |

**LANGUAGE REFERANCE:**

|  |  |
| --- | --- |
| **Language** | **Code** |
| English | en |
| Tamil | ta |
| Telugu | te |
| Malayalam | ml |
| Korean | ko |
| Chinese (Simplified) | zh-CN |
| Arabic | ar |
| Spanish | es |
| French | fr |

KEY ASSUMPTIONS:

**Voice Input** (Speech Recognition):

* Not all browsers support all languages for speech recognition.
* Chrome on mobile devices generally has the best support.

**Voice Output** (Speech Synthesis):

* Voice availability varies; use speech Synthesis.getVoices () to check available voices in your browser.

**Translation API**:

* This demo uses a free API (mymemory.translated.net) which may throttle or limit requests.
* For professional use, switch to the **Google Cloud Translate API**.

EXPECTED OUTCOME:

* Support for **multiple languages**, including English, Tamil, Telugu, Malayalam, Korean, Chinese (Simplified), Arabic, Spanish, and French.
* **Automatic language translation** of the spoken text using a public or commercial translation API.
* **Speech synthesis output** of the translated text in the target language using the browser's Speech Synthesis API.

**CONCLUSION:**

In conclusion, our Language Translation via Speech project demonstrates the practical integration of speech recognition, language processing, and speech synthesis to bridge communication gaps across different languages. By converting spoken input into text, translating it into the desired language, and then vocalizing the translated output, the system enables real-time, user-friendly multilingual interaction. This technology holds significant promise in areas such as travel, education, business, and emergency response, where language barriers often hinder effective communication.