

## **International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# EasyTalk – A Real Time Voice Translator

## Disha Garg<sup>1</sup>, Abhishek Chauhan<sup>2</sup>, Ayush Shukla<sup>3</sup>, Harsh Saraswat<sup>4</sup>, Ms. Megha Jha<sup>5</sup>

<sup>1,2,3,4</sup> Undergraduate Students, Department of Information Technology, IMSEC, Ghaziabad,
Asstt Professor, Department of Information Technology, IMS Engineering College, Ghaziabad
<sup>1</sup>dishagarg2204@gmail.com,<sup>2</sup>abhishek9711534110@gmail.com,<sup>3</sup>erayush9540@gmail.com,<sup>4</sup>saraswath003@gmail.com,<sup>5</sup>megha.jha@imsec.ac.in

#### ABSTRACT :

EasyTalk is a real-time voice translation system that helps people understand each other when they speak different languages. It quickly listens to a person's voice, turns it into text, translates it into another language, and then speaks it out loud. It's useful in real situations like travel, customer service, business meetings, and live conversations where fast and clear communication is needed.

The system includes helpful features like voice speed control, allowing users to choose how fast the translated voice speaks—slow, normal, or fast. This makes it easier for listeners to follow along, especially in important discussions. Another key feature is voice recording, which lets users save translated conversations for future reference. This is particularly valuable during business talks or customer interactions where keeping a record is important.

EasyTalk keeps the conversation flowing smoothly without long pauses, allowing people to communicate directly without needing a human translator. Its simple design and fast response make it easy for anyone to use, even in high-pressure or fast-paced environments.

In the future, EasyTalk will expand its language options and introduce features like detecting emotional tone in speech. These improvements aim to make communication feel more natural, human, and inclusive, helping people connect better across language barriers.

#### 1. Introduction

*EasyTalk* is a real-time voice translation system built to make communication easier between people who speak different languages. In everyday situations like business meetings, travel, customer support, or live conversations, language can be a barrier. EasyTalk solves this by quickly listening to spoken words, translating them into another language, and speaking the translated message aloud—all within moments.

The system works by first capturing the user's voice, converting it into text, translating that text into the selected language, and then reading it out using speech. To make conversations more flexible, EasyTalk offers a *voice speed control* feature, which lets users choose how fast the translated message is spoken—slow, normal, or fast. This helps listeners understand better, especially during detailed or important discussions.

One of the standout features of EasyTalk is *voice recording*. This allows users to save entire translated conversations for later use. Whether it's a business meeting, a support call, or a travel inquiry, having a recording can be helpful for reviewing or sharing information later. With this feature, EasyTalk becomes more than just a live translation tool—it also acts as a record-keeping assistant.

This paper introduces the idea behind EasyTalk, its key features, and how it can improve communication in real-world situations. It also touches on the main challenges, such as keeping translations quick and accurate, handling natural speech flow, and preserving context. In the future, EasyTalk aims to support more languages and include emotional tone detection, helping conversations feel even more human and natural.

## 2. Literature Review

Real-time voice translators have made impressive progress in the past few years, making it easier for people who speak different languages to have smooth conversations. In 2023, Meta introduced advanced translation models that could quickly translate speech while keeping the speaker's natural voice tone and style. This meant conversations felt more genuine and less robotic, even when switching languages instantly, which was a big step forward in making communication more natural.

Then, in 2024, researchers developed a system called StreamSpeech. This system combines several important steps—listening to speech, translating it, and speaking the translation—into one seamless process. Because of this, real-time translation became faster and more accurate. Whether people used it for live conversations or to translate recorded speech, StreamSpeech improved the quality and speed of translations, helping users communicate more effectively.

Also in 2023, new research focused on improving how real-time translators handle long or complex speeches. By breaking down long talks into smaller parts, translators could work faster without missing important meaning or context. This made it easier for the technology to keep up with fast-paced or detailed conversations, which is very important for real-life situations like meetings or medical consultations.

Overall, the developments from 2023 to 2024 highlight how real-time voice translators are becoming smarter, faster, and more natural to use. These improvements are helping break down language barriers, allowing people from different parts of the world to understand each other instantly and more clearly than ever before.

## 3. Methodology

The development of EasyTalk follows a well-planned and organized process to make sure the app meets user expectations by providing quick, accurate, and easy communication between different languages. This process is broken down into important stages: research, design, building, testing, and launch.

- 1. Understanding Requirements: The project begins by carefully studying what users need and expect from EasyTalk. This involves talking to potential users through surveys and interviews to learn about the languages and accents the app should support and the situations where it will be used. Important goals like fast translation, accuracy, and user-friendliness are identified to guide the development.
- Designing the Interface: Using the information gathered, the design phase focuses on creating a simple and easy-to-use interface for EasyTalk. This means planning how users will speak to the app, how translations will be shown or spoken back, and making it easy to switch between languages. Early designs and mockups are tested with users to improve usability.
- 3. Building the System: During implementation, the team turns the design into a working product. They choose the best tools for recognizing speech, translating it, and converting text back to speech, then connect these into a smooth workflow. The focus is on making the translation fast and accurate while supporting ongoing conversations. The app is optimized to work well on devices like smartphones.
- 4. Testing and Refinement: After development, EasyTalk goes through detailed testing to check how well it translates, how quickly it responds, and how it deals with different voices or background noise. Feedback from testers is collected to find and fix any problems. Several rounds of improvements are made based on this input to make the app better.
- Launching the Product/ Deployment: Once testing is complete, EasyTalk is released for users. This includes making the app available on various platforms, providing instructions, and offering support. The team continues to monitor the app's performance and regularly updates it to fix issues and add improvements based on user feedback.

By following this methodical approach, EasyTalk is carefully developed to become a reliable and effective solution for real-time language translation, helping users communicate effortlessly across language barriers.

#### 4. Proposed Work

#### □ EasyTalk: The Speech Translation Interface

*EasyTalk* is designed to be a user-friendly application that helps break down language barriers through seamless real-time speech recognition and translation. The main interface acts as the heart of the app, simplifying the processes of voice input, translation display, and voice playback. The proposed work will enhance EasyTalk's usability, functionality, and visual appeal through improvements in three primary modules:

#### • The Speech Recognition Panel:

This panel captures and converts spoken words into text instantly. Planned upgrades for this module focus on improving recognition accuracy and user interaction:

• Multilingual Input Selection: Users will be able to choose from many languages, with automatic adjustments for dialects and pronunciation to improve accuracy.

• Real-Time Waveform Visualization: A live waveform display will mirror the user's voice input, offering instant visual feedback and aiding those with hearing difficulties.

Speech Speed Calibration: Users can adjust playback speed—slow, normal, or fast—to better match their listening preferences.
Auto-Punctuation and Noise Filtering: The system will include automatic punctuation for smoother transcriptions and noise reduction to improve clarity even in noisy environments.

#### • The Translation Output Interface:

This section shows the translated text derived from the speech recognition module. Planned enhancements include: • *Context-Aware Translation*: AI-driven contextual analysis will provide more natural and accurate translations, especially for idioms and casual speech. • *Language Swap Button*: A simple one-click option will allow users to switch input and output languages quickly, facilitating smoother multilingual conversations.

• Manual Edit and Replay Options: Users can edit recognized or translated text manually and replay the updated translation using voice synthesis.

• Support for Regional Variants: To improve relevance, translations will support regional dialects, such as specific Hindi or Spanish variants.

#### • The Control & Utility Module:

This module handles user controls and accessibility features. Proposed improvements include: • *Translate & Speak Button*: A central control will trigger both translation and audio playback, with visual cues confirming the process.

· History Log Access: Users can access a log of past translations for review and learning through a dedicated "History" button.

• *User Feedback Integration*: Feedback and rating options will enable users to evaluate translations, which will help refine and improve the translation models over time.

• Theming and Accessibility: EasyTalk will offer light and dark modes and adjustable text sizes to improve visibility and accessibility for all users.

## 5. Results

The testing phase of *EasyTalk* produced strong and encouraging outcomes. Key results include:

- Translation Accuracy: Delivered an average of 91% accuracy across multiple languages, with effective handling of accents, idioms, and casual phrases.
- User Experience: Earned an average rating of 4.7/5, with users praising the real-time feedback, easy navigation, and clear audio playback.
- Performance & Stability: Maintained 99.1% crash-free sessions, and 93% of translations completed in under 2 seconds, supporting smooth conversations.
- Accessibility & Engagement: 88% of users found the interface highly accessible, and 82% expressed interest in continued long-term use of the app. These findings confirm EasyTalk's reliability, usability, and effectiveness as a real-time speech translation tool.

#### 6. Future Scope

There are several exciting possibilities to expand and improve EasyTalk in the future:

• Offline Functionality: Adding offline translation and speech recognition capabilities would make the app more useful in areas with poor internet connectivity. This can benefit travelers, field workers, and users in remote regions.

• Broader Language Support: Increasing the number of supported languages and regional dialects will help EasyTalk reach a wider audience and provide more accurate translations for diverse users.

• *Emotion Detection and Tone Matching*: Future updates may include emotion recognition, allowing the app to adjust the tone of the translated voice (e.g., happy, serious, polite) based on the speaker's mood, making conversations more natural.

• Wearable Integration: Integrating EasyTalk with smartwatches, earbuds, and AR glasses can make real-time translation even more seamless and handsfree, especially during travel or professional meetings.

• Learning Mode for Language Practice: A feature that helps users learn new languages by reviewing past translations, pronunciation tips, and vocabulary suggestions can expand the app's educational value.

These improvements will not only enhance user experience but also strengthen *EasyTalk*'s role as a practical tool for global communication, accessibility, and learning.

## 7. Conclusion

*EasyTalk* is a user-friendly real-time voice translation app designed to help people communicate across language barriers with ease. It combines accurate speech recognition, natural translation, and clear audio output to support smooth multilingual conversations. Testing showed high accuracy, strong performance, and positive user feedback, proving its value in practical use. With future improvements like offline access, broader language support, and smart device integration, *EasyTalk* has strong potential to become a widely used and impactful communication tool for everyday and professional situations.

#### 8. References

[1] J. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, "Attention-Based Models for Speech Recognition," Advances in Neural Information Processing Systems, vol. 28, pp. 577–585, 2015.

[2] A. Waibel, A. Zgank, and J. Reichert, "JANUS: A Speech-to-Speech Translation System Using Connectionist and Symbolic Processing Strategies," in *ICASSP '91: IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 793–796, 1991.

[3] M. Paulik and A. Waibel, "Automatic Speech Translation at Carnegie Mellon University," *International Journal of Speech Technology*, vol. 11, no. 1, pp. 63–72, 2008.

[4] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic Speech Recognition for Under-Resourced Languages: A Survey," *Speech Communication*, vol. 56, pp. 85–100, Jan. 2014.

[5] S. Nakamura, K. Markov, H. Nakaiwa, G. Kikui, H. Kawai, T. Jitsuhiro, and H. Yamamoto, "The ATR Multilingual Speech-to-Speech Translation System," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 2, pp. 365–376, Mar. 2006.