IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

Real Time Speech to Sign Language Translation

Sneha Anil Jagmalani[1], Dhanashri Anant Petkar[2], Shruti Rajesh Waychal[3], Revati Ramesh Salokhe[4], Prof. Dr. Ganesh V. Patil[5]
Students, CSE (Data science)[1,2,3,4]
Head of the Department, CSE (Data Science)[5]
DY Patil College of Engineering & Technology, Kolhapur, India

Abstract: *The project "Real Time Speech to Sign Language Translation" focuses on building a communication bridge between hearing individuals and the speech or hearing-impaired community by translating spoken language into sign language in real-time. It promotes inclusivity and accessibility by allowing seamless interaction and understanding across different modes of communication. Key features of the system include real-time speech detection, dynamic translation into accurate sign gestures, and user-friendly visual outputs to represent sign language clearly. The system is designed to handle basic to moderately complex sentences and can be used in various day-to-day settings like classrooms, public places, and workplaces. Overall, the project empowers the hearing-impaired community while encouraging empathy and awareness in society.*

Keywords: Hearing-impaired, real-time speech translation, American Sign Language, Speech Recognition, NLP, Sentiment Analysis, ASL Gestures, Emotion Detection

I. INTRODUCTION

In a world where communication is a fundamental aspect of human interaction, it is essential to ensure that everyone has equal access to effective communication methods, regardless of any barriers they may face. One of the most prominent challenges faced by individuals with hearing impairments is the lack of accessibility to real-time spoken communication. While traditional sign language interpreters provide a valuable service, there is a growing need for innovative technologies to bridge the communication gap without relying solely on human intermediaries. The Real-Time Speech to Sign Language Translation project aims to address this need by creating an intelligent, web-based platform that empowers hearing-impaired individuals to engage in live conversations with ease and confidence.

This project goes beyond basic translation by introducing features that enhance the emotional and visual quality of communication. It not only ensures real-time responsiveness but also focuses on delivering a natural, intuitive user experience. The aim is to improve the daily lives of individuals with hearing impairments, allowing them to participate more actively in social and professional settings. Ultimately, the project contributes to building a more inclusive digital environment and serves as a stepping stone for future innovations in the field of accessibility and assistive technology.

Need of the work

According to the World Health Organization (2023), over 430 million people globally suffer from disabling hearing loss, and this number is expected to rise to 700 million by 2050 [11]. In the United States alone, about 11 million individuals identify as deaf or significantly hard of hearing. These statistics clearly highlight the urgent need for inclusive communication solutions. Our project aims to address this need by targeting this large and often underserved population, helping to reduce the communication barrier between hearing and non-hearing individuals through real-time speech-to-sign language translation.

The project is structured to be completed over a timeline of 5 months, starting with research and requirement analysis, followed by system development, testing, deployment, and continuous updates. It is designed to be user-friendly, responsive, and adaptable for daily use in settings such as schools, public places, and workplaces. By focusing on both accuracy and accessibility, the project contributes toward creating a more empathetic and inclusive society. You can

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

explore more insights through this blog post on Hand Talk Translator App and Wired's article on AI-powered sign language. [14].

Existing System

Existing sign language translation systems commonly follow approaches like Direct Translation, Statistical Machine Translation, or Transfer-based Architecture. Most ASL systems focus on text-to-sign video conversion or avatar animations. However, few support real-time speech input, especially within browsers. Many of them also require complex installations or are restricted to desktop applications. Our work stands out by providing a lightweight, browser-based system that translates spoken English into ASL using live speech recognition and dynamic sign rendering.

Proposed System

The proposed system is designed for individuals with hearing impairments who use ASL as their primary language. It addresses the communication barrier between hearing individuals and the deaf community by providing a real-time speech-to-ASL translation platform. Developed using JavaScript and modern front-end libraries, the system captures spoken English through the SpeechRecognition Web API, processes it using NLP techniques with support from external Gemini AI APIs for punctuation correction and sentiment analysis, and then translates the content into ASL gestures. These gestures are animated using Framer Motion and displayed using dynamically loaded sign images via Webpack's require.context. It also performs audio feature analysis—extracting frequency, amplitude, and spectral characteristics—visualized through Recharts graphs. This system is valuable for ASL learners, educators, or anyone aiming to improve communication with the ASL community. By combining accessibility, usability, and emotional context detection, the system enables more inclusive and meaningful interactions.

II. LITERATURE SURVEY

American Sign Language (ASL) has been the focus of numerous technological advancements aimed at facilitating communication for the deaf and hard-of-hearing community. One notable development is DeepASL, a system that utilizes infrared sensors to non-intrusively capture ASL signs. It employs a hierarchical bidirectional deep recurrent neural network (HB-RNN) alongside a Connectionist Temporal Classification (CTC) framework to enable both word-level and sentence-level translation. Evaluations of DeepASL demonstrated an average word-level translation accuracy of 94.5% and an 8.2% word error rate for unseen sentences, highlighting its potential in real-world applications [1].

In addition to vision-based systems, wearable technologies have been explored for ASL translation. The Myo armband, equipped with electromyography (EMG) sensors and inertial measurement units, has been utilized to detect muscle activity and motion for gesture recognition. Studies have shown that combining the Myo armband with machine learning algorithms, such as support vector machines (SVMs), can effectively classify a range of ASL gestures with high accuracy. This approach offers a portable and user-friendly solution for real-time sign language translation [2].

Further research has integrated the Myo armband with additional sensors to enhance ASL recognition capabilities. For instance, combining the Myo armband with devices like the Leap Motion controller allows for the capture of both muscular and spatial data, improving the system's ability to interpret complex gestures. Such multimodal systems have demonstrated improved accuracy and robustness in various environmental conditions, making them suitable for diverse applications [3].

Despite these advancements, challenges remain in achieving seamless and natural ASL translation. Systems must account for the nuances of ASL grammar and the variability in individual signing styles. Ongoing research aims to develop more sophisticated models that can adapt to these variations, ensuring accurate and fluid translations. The integration of advanced sensors, machine learning algorithms, and user-centric design continues to drive progress in this field [4].

Recent advancements in sign language translation have increasingly focused on improving the efficiency and accuracy of multimodal models. TSPNet (Temporal Semantic Pyramid Network), introduced by Li et al., leverages a hierarchical feature learning strategy for sign language translation. By modeling long-term temporal dependencies and semantic structures using a Temporal Semantic Pyramid, the model achieves improved accuracy in both isolated and continuous

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

sign language recognition. Evaluations on large datasets such as RWTH-PHOENIX-Weather show TSPNet's superior performance over traditional CNN-RNN pipelines, marking a significant advancement in temporal modeling for sign language systems [5].

Expanding the scope of sign language translation across cultures, Kumar et al. proposed a hybrid framework integrating large language models (LLMs) for bi-directional translation between American Sign Language (ASL) and Indian Sign Language (ISL). Their work emphasizes the creation of a parallel ASL-ISL dataset and the use of advanced transformer-based models to handle linguistic structure differences. The system shows promise in multilingual accessibility and showcases how deep learning can bridge regional sign language disparities. This approach opens possibilities for the development of globally adaptive sign language tools [6].

In a comprehensive review, Shahin and Ismail present a taxonomy and performance evaluation of sign language translation systems ranging from rule-based models to modern deep learning and transformer architectures. Their survey emphasizes the role of Natural Language Processing (NLP), Computer Vision, and Recurrent Neural Networks (RNNs) in sign language translation. The paper evaluates challenges such as limited datasets, signer variability, and grammar modeling, offering insights for future research directions. The survey serves as a foundational reference for understanding the evolution and capabilities of current ASL systems [7].

Chen et al. introduced a simple and effective multi-modality transfer learning baseline for sign language translation, showing how visual, spatial, and textual cues can be combined with minimal complexity. Their model outperformed previous baselines by leveraging pretrained features from different modalities and combining them using a simple transformer-based architecture. This proves that transfer learning and lightweight models can be both effective and scalable for real-time sign translation systems [8].

Earlier systems, such as that by Gunes and Piccardi, focused on speech-to-sign language translation, integrating spoken language recognition and visual output generation. Their platform translated spoken English into signs by combining speech recognition with rule-based mapping to sign sequences. Though limited by grammar handling and gesture animation fidelity, this system marked an early step toward real-time translation for deaf users and paved the way for deeper AI integration in future systems [9].

Finally, Cheok et al. designed a sign language translation system using Convolutional Neural Networks (CNNs) trained on image frames of sign gestures. Their research showed that even with a basic CNN architecture, accurate classification of isolated signs is achievable, especially with controlled backgrounds. Their study encourages the use of computer vision as a lightweight yet powerful approach for gesture-based translation, especially in resource-constrained environments [10].

III. MOTIVATION

The motivation behind developing a Real-Time Speech to Sign Language Translation system stems from the need to bridge the communication gap between hearing individuals and the deaf or hard-of-hearing community. In everyday life, individuals with hearing impairments often struggle to follow spoken conversations due to the lack of accessible, real-time translation solutions. This project aims to make verbal communication instantly understandable by converting speech into sign language in a seamless and interactive way. The broader goal is to support a more inclusive society where communication is equitable across all spaces—especially in public services, education, and healthcare.

Beyond its practical use, the system introduces emotional intelligence by considering the tone and sentiment behind spoken words, making interactions feel more natural and human-like. These emotionally-aware features enhance user experience and foster a deeper sense of connection. The solution holds potential to scale into different domains such as mobile platforms, digital learning environments, or service kiosks. It also encourages greater awareness and sensitivity towards the challenges faced by the deaf community, promoting empathy and inclusive design principles. Ultimately, the project highlights how technology can play a transformative role in building an accessible and equitable world where communication is a universal right, not a privilege.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

IV. METHODOLOGY

1. Input Audio :

The system captures real-time speech from the user through a microphone using the Web Speech API or a similar voice input tool.

2. Feature Extraction :

Audio features such as pitch, frequency, and amplitude are extracted using the Web Audio API (AudioContext). These features help in understanding the nature of the speech (e.g., emphasis, stress, etc.).

3. Learning Environment :

The extracted features and recognized text are passed into a structured environment where predefined mappings, models, and logic guide the translation process.

4. Language Model :

A lightweight language model interprets the sentence structure and helps correct or complete partial speech data. This improves the contextual understanding of the sentence.

5. Speech Recognition :

The live audio input is converted into text using real-time speech recognition (e.g., Web Speech API). The output is a readable sentence or phrase.

6. Recognized Word Processing :

The recognized text is tokenized to extract important keywords that are directly mapped to sign language gestures or animations.

7. Show Related Animation :

Based on the identified keywords, corresponding sign language animations (GIFs or vector motions) are loaded and displayed smoothly using libraries like framer-motion.

8. Sentiment Analysis :

The full sentence is analyzed for overall sentiment (positive, negative, or neutral) using basic NLP sentiment models. This gives emotional context to the speech.

9. Emotion Detection :

A deeper emotion detection process identifies specific emotions like happiness, sadness, anger, etc., and adjusts the animation tone (e.g., facial expressions or speed) accordingly for a more natural output.

V. SYSTEM DESIGN & IMPLEMENTATION

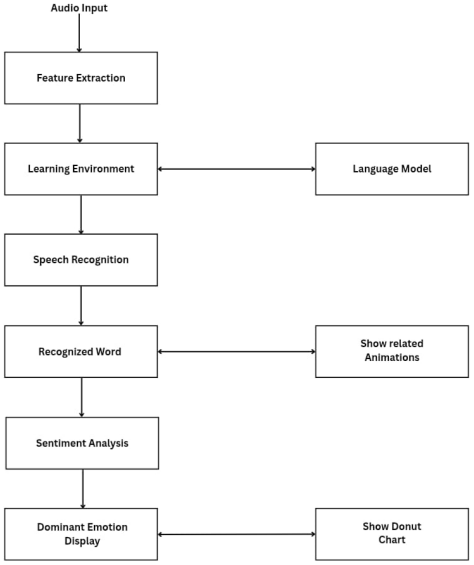


Fig 1. System Architecture of Real Time Speech to Sign Language Translation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

1. User Input Layer (Microphone Input)

o The user speaks into the device's microphone, which acts as the primary input to the system.

o The microphone captures real-time audio and sends it to both the speech recognition engine and the audio analysis module simultaneously.

2. Speech Recognition Module (Web Speech API)

o This module receives the audio input and converts it into text using the SpeechRecognition interface from the Web Speech API.

o It processes the speech in real-time, generating word-by-word transcripts.

o The output of this module is plain text, which is passed to the sign mapping module.

3. Audio Analysis Module (Web Audio API)

o In parallel, the audio input is processed by the AudioContext of the Web Audio API.

o This module extracts audio features like frequency, amplitude, and spectral patterns.

o These features are sent to the visualization module to generate real-time graphs.

4. Text-to-Sign Mapping Module

o This module takes the transcribed text and matches each word to a corresponding sign language image or symbol.

o It uses require.context() in JavaScript to dynamically load and retrieve image assets based on the spoken words.

o If a word has no available sign, a default icon or message is displayed.

5. Sign Rendering and Animation Module

o Once the sign images are selected, they are passed to this module, which handles the visual display.

o Using Framer Motion, it animates the transition between one sign and the next for a smooth visual experience.

o This helps mimic the natural flow of human sign language gestures.

6. Visualization Layer (Charts and Audio Feedback)

o The audio features from the analysis module are visualized using Recharts, a React-based charting library.

o Line charts or bar graphs display amplitude, frequency, and speech energy in real-time.

o This layer enhances interactivity and provides visual feedback to the user.

7. Frontend Interface (User Interface Layer)

o Built using React.js, this layer brings all components together into one cohesive interface.

o It includes:

• A start/stop microphone button

• A text display area for live transcription

• A sign language display panel

• An audio chart section

o It handles user interaction, real-time updates, and makes the application responsive across devices.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

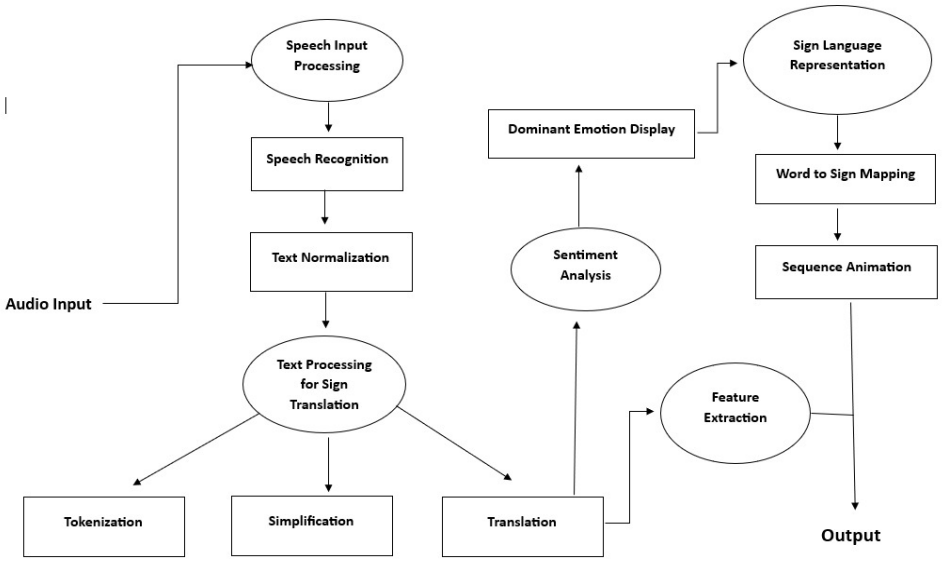


Fig 2. Implementation Diagram

Implementation Details:

1. Speech Recognition Using Web Speech API

The project uses the Web Speech API, specifically the SpeechRecognition interface, to convert real-time spoken words into text. When the user speaks into the microphone, the API captures the audio and processes it to generate a text transcript. This happens continuously so the system can keep updating the spoken content. It handles voice input smoothly and works well in browsers like Google Chrome and Microsoft Edge. The recognized text is then sent to the sign mapping module for further processing.

2. Audio Feature Extraction with Web Audio API

In addition to speech-to-text conversion, the system also analyzes the audio signal using the Web Audio API, particularly through the AudioContext interface. This module helps extract low-level audio features such as:

- Frequency (pitch)
- Amplitude (volume)
- Spectral energy (distribution of energy over frequency)

These features provide insight into the nature of the spoken input. Although not directly used for translation, they are displayed as visual feedback to users for better interactivity.

3. Real-Time Visualization Using Recharts

The visual representation of audio features is handled using the Recharts library, a popular charting library for React. As the user speaks, different graphs such as line charts or bar graphs are updated in real-time to show:

- The waveform of the audio input
- Frequency variations
- Changes in amplitude

This feature makes the application visually informative and adds an analytical aspect for the user.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

4. Mapping Speech to Sign Language Icons

Once the speech is converted into text, each word is matched with a corresponding sign language symbol or image. These signs are stored as static image files (e.g., PNG or SVG). The mapping is performed using JavaScript's require.context() function, which dynamically loads image files based on the spoken words. This allows the app to scale easily and manage a large number of sign images efficiently.

5. Smooth Animation with Framer Motion

To enhance the user experience, the transition between signs is animated using Framer Motion, a powerful animation library for React. When a new word is spoken, its corresponding sign image appears with smooth motion effects, such as fade-in or slide transitions. This makes the interaction more natural and engaging.

6. Clean Visual Elements with Lucide-react

Icons and visual cues in the application are displayed using Lucide-react, a modern icon library. These icons are used for UI components like the microphone button, audio status indicators, and user notifications. It keeps the interface visually clean and user-friendly.

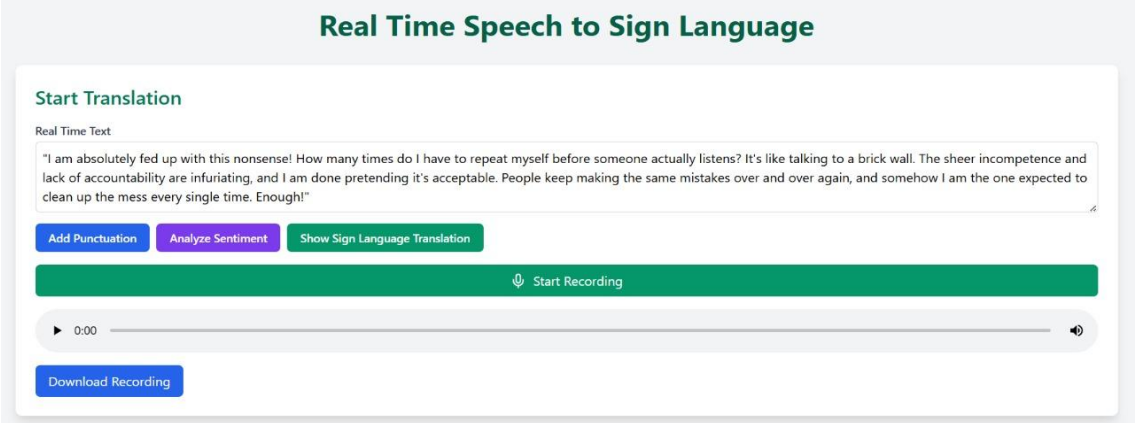
7. Frontend Development Using React.js

The entire user interface of the project is built using React.js, a JavaScript library for building component-based UIs. The interface includes:

- A microphone control button to start and stop speech recognition.
- A text area that displays real-time transcriptions of spoken words.
- A sign display panel that shows corresponding sign language images or icons.
- A visual chart area where audio features are plotted in real-time.

The application is fully responsive and can run on both desktop and mobile browsers, making it accessible to a wide range of users.

VI. RESULT ANALYSIS



This image showcases the input interface where users can provide audio input in two ways :

o The "Start Recording" button enables users to record their voice in real time.

o Alternatively, users can manually type the text into the "Real Time Text" input box.

This dual-mode input flexibility ensures better accessibility and ease of use depending on the user's preference.

Additionally, three functional buttons have been added to enhance user control:

o "Add Punctuation" button processes the raw text using the Gemini API to add punctuation and capitalization.

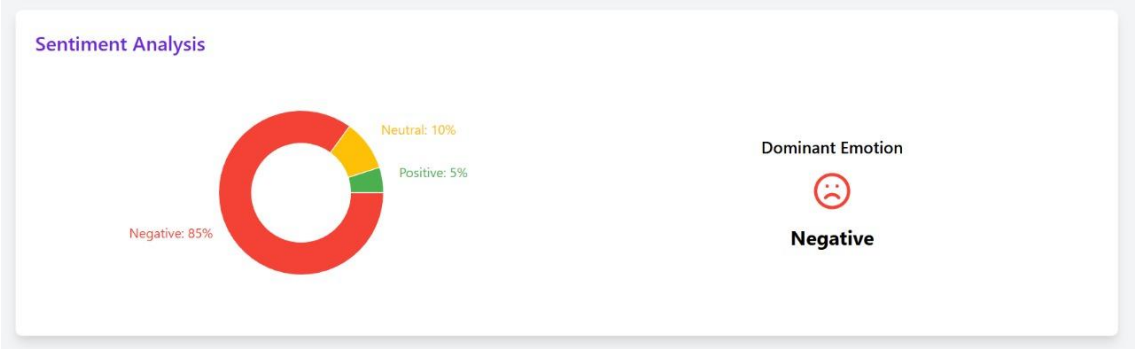
o "Analyze Sentiment" button detects the emotional tone of the text and displays sentiment scores.

o "Show Sign Language Translation" button converts the processed text into animated ASL signs.

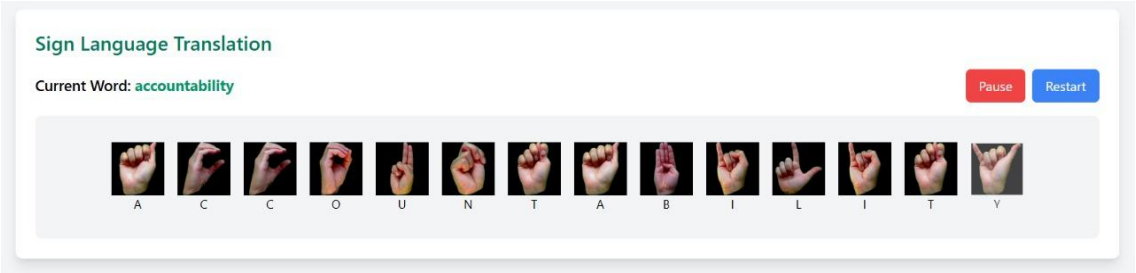
DOI: 10.48175/IJARSCT-26488

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

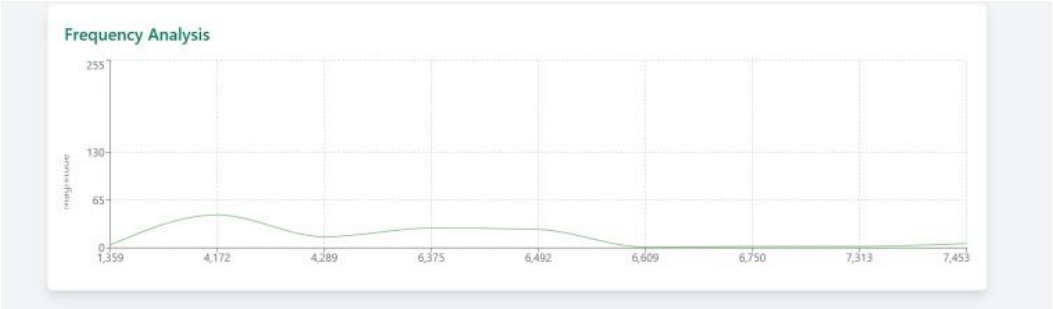
Users can also:
o Listen to the Recorded Voice by clicking the play button after recording.
o Download the Audio by using the provided "Download Recording" button to save the recorded file.



This image showcases the sentiment analysis features of the application :
o The "Analyze Sentiment" button detects the emotional tone of the transcribed text and categorizes it into positive, negative, or neutral.
o On the left side, a donut chart visually represents the sentiment distribution with three categories :
• Positive (green)
• Negative (red)
• Neutral (yellow)
o On the right side, the dominant emotion is displayed using a simple emoji icon, representing the overall emotion detected in the text.
This analysis provides users with a clear, visual representation of the emotional context of the speech.

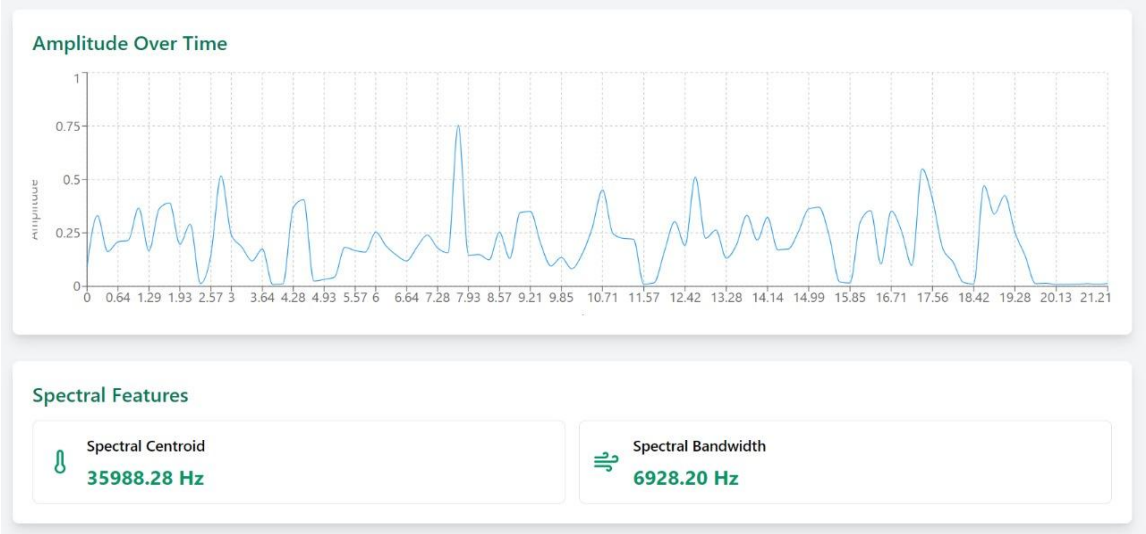


This image displays an animation of sign language gestures corresponding to the text entered or spoken:
o Each word from the input is translated into a visual sign language representation letter by letter.
o The animation shows one word at a time, allowing users to see the complete translation before moving to the next word.
o This mechanism helps users visually understand how each part of the text is signed, creating an intuitive connection between language and signs.



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

This image provides a detailed frequency spectrum analysis of the recorded voice.
o It shows how different frequencies are distributed in the audio signal.
o This analysis is crucial for understanding pitch, tone, and overall audio quality, and can be used in further signal processing or classification tasks.



This image presents both spectral features and time-based audio analysis.
o At the top, a waveform graph shows the amplitude over time, illustrating how loudness varies as the audio progresses.
o Below that, it extracts spectral centroid and spectral bandwidth, which are key features in analyzing the tone and clarity of the audio.

VII. CONCLUSION

The Real-Time Speech to Sign Language Translation project successfully demonstrates the potential of assistive technology in bridging communication gaps between the hearing and speech-impaired community and the rest of society. By converting spoken language into corresponding sign language gestures in real-time, this system offers an accessible and inclusive solution for everyday conversations.

1. Functionality and Efficiency

- The application effectively captures live speech using the Web Speech API and processes it using audio analysis techniques like frequency, amplitude, and spectral features.
- The visual feedback is delivered through intuitive animations and icons, built using React, Lucide-react, and Framer Motion, ensuring a responsive user interface.

2. Innovative Use of Frontend Technologies

- Tools like Recharts provided visual insights into speech patterns, enhancing user understanding and system transparency.
- The use of require.context and modular code structure allowed easy management of sign assets and real-time translation logic.

3. Impact and Social Relevance

- The system promotes inclusivity, accessibility, and empowerment for the speech- and hearing-impaired community.
- It can be expanded into various domains like education, customer service, healthcare, and public service announcements.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

VIII. FUTURE SCOPE

1. Multilingual and Region-Focused Expansion

- A key future enhancement involves extending the system to support multiple sign languages based on regional and national variations, such as Indian Sign Language (ISL), British Sign Language (BSL), etc.
- This would require creating a robust dataset from scratch, tailored to the grammar, structure, and gesture
- nuances of each specific sign language.
- Users would be able to easily switch between languages based on their preference or region, and the application interface itself could dynamically change to reflect the selected spoken or sign language.

2. Bidirectional Communication

- Incorporating gesture recognition through computer vision techniques would allow sign-to-speech
- translation, making the system fully interactive.
- This would enable a seamless two-way conversation between hearing and hearing-impaired individuals.

3. Mobile and Offline Accessibility

- The system promotes inclusivity, accessibility, and empowerment for the speech- and hearing-impaired community.
- It can be expanded into various domains like education, customer service, healthcare, and public service announcements.
- Future versions could be deployed as cross-platform mobile applications to provide on-the-go support.
- Offline capabilities would enhance usability in low-internet or rural areas.

4. Challenging Data Science Opportunity

- Building region-specific and multilingual sign language models is a huge and complex task, involving challenges in data collection, annotation, gesture modelling, and contextual translation.
- This presents a significant research opportunity in the Data Science domain, involving natural language understanding, gesture synthesis, transfer learning, and scalable deployment.
- The work will not only push the boundaries of accessible technology but also test the limits of what data science can achieve in human-centric problem-solving.

REFERENCES

[1]. Biyi Fang, Jillian Co, Mi Zhang; "DeepASL: Enabling Ubiquitous and Non-Intrusive Word and Sentence-Level Sign Language Translation", Proceedings of the 15th ACM Conference on Embedded Networked Sensor Systems (SenSys), DOI: 10.1145/3131672.3131693, November 2017.

[2]. He, J., & Yang, Y.; "Hand Gesture Recognition Using MYO Armband", 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), DOI: 10.1109/ICCSEC.2017.8446906, December 2017.

[3]. Hussein Naeem Hasan; "A Cost-Effective Deaf-mute Electronic Assistant System Using Myo Armband and Smartphone", arXiv preprint arXiv:2503.14901, DOI: 10.48550/arXiv.2503.14901, March 2025.

[4]. Hussein F. Hassan, Sadiq J. Abou-Loukh, Ibraheem Kasim Ibraheem; "Teleoperated Robotic Arm Movement Using EMG Signal with Wearable MYO Armband", arXiv preprint arXiv:1810.09929, DOI: 10.48550/arXiv.1810.09929, October 2018.

[5]. Dongxu Li, Chenchen Xu, Xin Yu, Kaihao Zhang, Ben Swift, Hanna Suominen, Hongdong Li; "TSPNet: Hierarchical Feature Learning via Temporal Semantic Pyramid for Sign Language Translation", arXiv preprint arXiv:2010.05468, October 2020.

[6]. Malay Kumar, S. Sarvajit Visagan, Tanish Sarang Mahajan, Anisha Natarajan; "Enhanced Sign Language Translation between American Sign Language (ASL) and Indian Sign Language (ISL) Using LLMs", arXiv preprint arXiv:2411.12685, November 2024.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, May 2025
Impact Factor: 7.67

[7]. Nada Shahin, Leila Ismail; "From Rule-Based Models to Deep Learning Transformers Architectures for Natural Language Processing and Sign Language Translation Systems: Survey, Taxonomy and Performance Evaluation", arXiv preprint arXiv:2408.14825, August 2024.
[8]. Yutong Chen, Fangyun Wei, Xiao Sun, Zhirong Wu, Stephen Lin; "A Simple Multi-Modality Transfer Learning Baseline for Sign Language Translation", arXiv preprint arXiv:2203.04287, March 2022.
[9]. H. Gunes, M. Piccardi; "Spoken Language to Sign Language Translation System Based on Speech Recognition", Proceedings of the 11th International Conference on Multimodal Interfaces, DOI: 10.1145/3364908.3365300, November 2019.\
[10]. M. J. Cheok, Z. Omar, M. H. Jawar; "Sign Language Translation System Based on CNN Model", International Research Journal of Engineering and Technology (IRJET), Volume 11, Issue 5, May 2024.
[11]. WHO (2023): https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss
[12]. National Deaf Center (2021): https://nationaldeafcenter.org/faq/how-many-deaf-people-live-in-the-united-states
[13]. Wired (2025): https://www.wired.com/story/silence-speaks-deaf-ai-signing
[14]. Hand Talk Blog: https://www.handtalk.me/en/blog/meet-the-hand-talk-sign-language-translator-app/