# Speak-to-Sign: A Real-Time NLP-Driven Framework for Indian Sign Language Translation : A Review

Mrs S.Sudeshna[1], B Sai Nikhitha[2], B Vinaya[3], D.Dhanasree[4], D Chandrika[5]

[12345] CSE Department, VNRVJIET

Sudeshna611@gmail.com;sainikhitha695@gmail.com;vinayaboinpally@gmail.com;dhanasreedaidi@gmail.com;chandrikad2004@gmail.com;

**Abstract.** The primary impact of communication barriers on the lives of many deaf or hard-of-hearing individuals lies in their ability to understand and comprehend most spoken languages. Accordingly, this paper presents an effective system which can convert speech to Indian Sign Language (ISL) via images and gifs. The proposed system utilizes Python with Google's Speech Recognizer API contact to convert speech to text and then applies NLP on the text to generate ISL tokens that are rendered as gestures for user-friendly access to the deaf. The major features include efficient speech-to-text conversion and NLP-driven accurate processing and gesture generation in ISL to augment the efficiency of existing solutions in translation, processing speed, and privacy issues. The system promotes flexibility via the most current algorithms and a modular design to be easily added features or modified in the future. It enhances access to education, employment, and social participation for the hearing impaired and proclaimed as a contribution toward a better society. This offers more than just an aid for communication systems; it promotes inclusivity and technology itself among innovations.

**Keywords:** Audio-to-Sign Language Conversion, Indian Sign Language (ISL), Natural Language Processing (NLP), real-time processing, sign language visualization, gesture-based communication, accessibility solutions, inclusive



S. K. Swarnkar and Y. K. Rathore (eds.), *Proceedings of the International Conference on Advances and Applications in Artificial Intelligence (ICAAAI 2025)*, Advances in Intelligent Systems Research 194,
https://doi.org/10.2991/978-94-6463-738-0_30

technology, linguistic tokenization, multimodal interaction, assistive communication tools, dynamic gesture generation, hearing-impaired accessibility, Human-Computer Interaction (HCI), scalable assistive systems.

## 1 Introduction

Communication is the back of human interaction. It may allow people to share ideas, emotions, and knowledge necessary for building cooperative teams, social interaction, and individual improvement. Millions of people all over the world, particularly the deaf or the hard of hearing, still consider effective communication to be a big challenge. With such a dearth of available tools and technologies that could be used to translate the spoken word to a modality intelligible by people who have a hearing problem, people become marooned during those events of life, including education, work, and socialization. Estimates of the WHO claim that over 5% of the global population is suffering from hearing loss in some degree, and for these individuals this often represents their best opportunity to express themselves using signs. This cannot be the same in India, as for obvious reasons of linguistic diversity things become much complex. Even with Indian Sign Language, there is no implementation and incorporation into the media channels. This alienates the deaf society. Solutions employing these technologies capable of using assistive features, for instance speech-to-text converter, gesture recognition system, sign language translators are coming forth; still solutions do not help the hearing impaired society of India in general. The present systems show poor real-time processing, limited vocabulary support and dependence on the non-scalable datasets. They are not culturally and linguistically specific to much of an extent, which is one of the major requirements in a country like India, where the regional languages and dialects are quite different. In order to deal with communication, the problem of barriers becomes even worse in relation to contextualizing to Indian Sign Language, thus effectively allowing individuals to participate in education, professional, and social forums. This calls up an imperative need to further provide more efficient and practical solutions for communication. Now, at the pace of NLP and machine learning technology as advanced as it is, it becomes opportune time to develop an audio-to-sign language translator that could help bridge communication gaps for the deaf. It provides a proof-of-concept Audio-to-Sign Language Translator prototype that transcribes given spoken languages into ISL gestures in real time. The input speech is used to process through advanced speech recognition algorithms, NLP techniques, and open-source tools like the Google Speech Recognizer API and NLTK libraries to convert those spoken words into text and translate that text further into ISL gestures represented as images or GIFs. Of importance in India, this facilitates improving communication, education, and social interaction among its hearing-impaired population. The adaptive variety in its response of the deaf community toward the needs gives systems-based design at real time process, lingual accuracy, and relevant to culture compared with conventional techniques which are based on not changing the data record or orthodox library of gestures put in rigid frames; this opposite method includes a dynamic function with NLP which techniques are followed at tokenization and represents complex structures presented where vocabulary may vary. It would have modular architecture, hence extending the system into the future, adding new languages, gestures, and features in order to maintain the relevance of the system over a more extensive period. Three objectives of this paper are to develop and test the Audio-to-Sign Language Translator prototype that efficiently translates spoken words into ISL gestures, to check how it performs in terms of its ability to take real-time capability, linguistic exactness, and adaptability toward Indian linguistic domains, and also to contribute to a growing body of knowledge about assistive technologies as a scalable solution that can grow to be linguistically and geographically extended later. This work overcomes the communication barriers of the hearing-impaired community, hence giving access for their participation in various sectors of life. Advanced speech recognition and NLP technologies are used based on the upcoming system to present a solution to the communication challenges faced by the hearing-impaired community in India. This would open avenues for research and development in the areas of speech-to-sign translation, human-computer interaction, and assistive technology development. It will form a more accessible society and inclusive society for everyone.

## 2 Related Work

In the paper Bi-directional translation methods of ASL and speech/text Amira Elnashar[1] proposed a bi-directional system that translates American Sign Language (ASL) and speech/text to help mute and hard-of-hearing people. Using data gloves, cameras, and machine learning, it achieves high letter-level accuracy and 77% word-level accuracy. Text or speech inputs are converted into animated ASL using 3D avatars. Plans also focus on augmentation of datasets and improving the accuracy.

Biplov Paneru, Bishwash Paneru, Khem Narayan Poudyal[2] proposed an AI-based system that translates ASL gestures into Nepali text and speech, thereby enhancing the communication of the hearing-impaired in Nepal. It uses ResNet50 and VGG16 for gesture recognition with high accuracy. The system has a Python-based GUI using OpenCV for real-time input and gTTS for Nepali audio output, thus making it more inclusive through advanced assistive technology.

Developing the multimodal dataset for the Arabic Sign Language production paper[3] introduced the new Arabic Sign Language (ArSL) Friday sermons, creating a more accessible resource for deaf people. There are audio recordings, text transcripts, and 262 video clips by expert signers that were annotated for correctly representing religious content. The testing ensures that it has a reliable translation, providing it to be of huge value in upcoming ArSL translation and recognition systems.

VoiceMitra[4] is a system that translates speech and text to sign language visuals in order to empower the deaf with communication. It offers Speech-to-Text, Language Translation, as well as Sign Language Conversion modules and comes with easy user interfaces to show translated phrases. VoiceMitra enhances the communication abilities of hearing-impaired individuals in education, daily life, and even entertainment.

Dr. J. Vijaya[5], and her team designed an audio-based Indian Sign Language (ISL) translation system based on MFCC feature extraction and GHMM speech-to-text conversion. The work done by NLTK in text processing provides videos in Indian Sign Language with a very high accuracy and ensures real-time performance thereby enhancing accessibility to hearing-impaired people.

Audio to Sign Language Converter with Python[6] discusses conversion of spoken audio into corresponding representations of sign languages to bridge communication gaps for people suffering from hearing problems. The real-time system using speech recognition techniques and mapping them onto sign languages is used here.

A speech to sign language translation for Indian languages [7] is the system that translated speech from six Indian languages into ISL using MFCC-GMM for speech recognition, LSTM for text translation, and direct mapping for ISL generation with a result of accuracy above 80%. Future work can include adding more dialects and 2D animation.

The Voice to Sign Language-A Review by G. Santhanakrishnan and S. Kumaravel [8] writes that they were deaf and dumb because they did not know sign language. This article discusses possible answers to these issues, which sums up the preceding mention of these technologies based on NLP and AI. Maybe the future trend will be AI advanced and in real time, for greater integration.

The Audio to Sign Language Translator Web Application[9] is a paper that introduces a web-based application to translate English text and voice into Indian Sign Language (ISL) using NLP. It uses a bilingual dictionary and ISL grammar rules to produce videos of ISL gestures, thereby enhancing communication for the hearing impaired. Though effective, it also mentions some of the limitations like a restricted vocabulary in ISL.

Identity Card services speech-to-sign language translation system for Spanish by R. San-Segundo et al.[10] translates the Spanish speech to LSE or Spanish Sign Language using an input speech recognizer system, a natural language translator and the playback call engine in 3D avatar format. Statistical translation outperformed the rule-based one with the results of the experimentations amounting to 31.6% Sign Error Rate and 0.578 score in BLEU. The study focuses on further reduction in translation latencies and improvement in real-time performance late-on using the eSIGN-enabled 3D avatar.

I.A. Adeyanju[11] critically reviewed the techniques of machine learning for automatic sign language recognition. It uses a total of 649 publications from 2001-2021 with the bibliometric methods that have identified the trends, techniques, and collaboration in ASLR research. It gives insights on improvements in the methods of vision-based recognition, data acquisition, and feature extraction but pointed out gaps in the development of intelligent ASLR systems. It has offered insights for the guidance of further research for creating robust, AI-powered sign language systems that enhance communication among the hearing-impaired.

M. M. Kamruzzaman[12] developed a system in which CNN is utilised to identify Arabic sign language letters and map them into Arabic speech. The model is trained through 31 hand signs with image preprocessing and augmentation, which gives the model a fair level of accuracy. It then gets connected to a text-to-speech engine, making the solution efficient to aid communication among hearing-impaired people.

As reported by Dheenesh Pubadi et al.[13], the ultimate purpose of Tamil speech-to-text code mixing and code switching is the development of a Tamil speech-to-text application that safeguards the language. That way, this avoids transcription with auto-correction, code-mixing, and code-switching apt to Tamil's particularities in phonetics. This renders the exact phonetic transcription and pronunciation to its natives and even to its non-natives. But it would not only achieve these objectives as it would act as an equipment for indigenous-designed learning, serving language preservation.

Mourad Broura and Abderrahim Benabbou[14] introduce ATLASLang MTS 1, a system translating Arabic text into Arabic Sign Language (ArSL) using rule-based and example-based methods. It uses tools such as SAFAR and ALKHALIL for analysis and generates ArSL via GIFs, 3D avatars, or finger spelling. The objective is to improve communication for the deaf, and future plans include expanding the database and enhancing 3D animations.

Triphone Modeling for Kannada Speech Recognition[15]: In this study, they have developed a Kannada speech recognizer using triphone modeling. The 39-dimensional MFCC features are fed into the system with tristate HMMs modeled by GMM. 46 phonemes (12 vowels, 34 consonants) are considered in this work along with their recognition performance as monoclass, word-internal triphone, and tied-state triphone models across various GMM configurations.

This effort, bearing the title: Voice and Speech Recognition in Tamil, is the creation of Kiran R. and others[16]. Their idea is to develop a speech recognition system designed for smartphones, which translates Tamil speech into text and creates different functionalities such as voice dialing or sending messages. They have accomplished this in a variety of ways, Dynamic Time Warping, Hidden Markov Models (HMM), Gaussian Mixture Models with support vector machines and neural networks as their machine learning approach. In training sessions, HMM achieved an accuracy of 100 percent, thus proving efficacy for Tamil speech recognition with astounding accuracy about 98 percent while testing.

The research effort titled "Audio to Sign Language Translation for Deaf People" by Ankita Harkude et al. [17] aims at developing a communication system that bridges the gap between normal and deaf people. This project translates audio inputs into Indian Sign Language (ISL) representations, making it easier to communicate. The system accepts the audio as an input, transforms it into text with the help of speech recognition by Google Speech API, and further processes the text using Natural Language Processing. Finally, it follows the grammar rules of ISL and shows the corresponding ISL gestures in terms of images or GIFs. The front-end design is done using EasyGUI, while PyAudio has been used to handle the audio input. This is an innovative system that enhances the communication and access to deaf people but does not integrate facial expressions yet, which are a significant aspect of sign language. The proposed future scope involves implementing this system for news channels, replacing human interpreters with automated sign language representation on screens and integrating facial expressions to enhance the overall experience.

Audio to Indian Sign Language Translator[18] This is a product designed by Harshita Mishra, Mansi Sharma, Muskan Ali Qureshi, and Shivani Chaudhary. This project proposal is to develop a system that can take audio as an input in Indian Sign Language (ISL). The system takes live speech or text inputs and uses techniques like NLP, removes stop words, extracts the root words, and maps these root words to the corresponding ISL videos in a pre-existing database. Finally, it generates a video sequence that demonstrates the ISL signs for the input given. It holds great promise in bridging the communication gap for India's deaf and hard-of-hearing population, with more potential in incorporating text-to-speech functionalities and expanding ISL vocabularies.

The research work titled Speech to Indian Sign Language Translator[19] by Mr. Deepak Bhagat, Mr. Hitesh Bharambe, Mr. Jashpal Joshi, and Prof. Shahe Gul [X] focuses on developing a system that bridges the communication gap for the deaf and mute communities. The system translates spoken language into Indian Sign Language (ISL) using tools like Google Speech Recognition, Natural Language Toolkit (NLTK), and TF-IDF. Speech input is converted into text, then into ISL grammar. Gestures will be represented by predefined images or GIFs. This model captures the temporal dependencies and utilizes more sophisticated machine learning algorithms such as Convolutional Neural Networks and Recurrent Neural Networks. It proved to achieve the most accuracy level in speech-to-sign translation that could serve to be an essential step for enhanced accessibility and inclusion for the deaf community.

This work, Empowering Sign Language Communication: Integrating Sentiment and Semantics for Facial Expression Synthesis [20], is the

work of Rafael V. Azevedo, Thiago M. Coutinho, João P. Ferreira, Thiago L. Gomes, and Erickson R. Nascimento. The goal of this project is to enhance sign language production by generating facial expressions that include sentiment and semantic information, which are essential for conveying grammar, emotions, and emphasis in sign language communication. It applies a graph convolutional neural network along with generative latent optimization in the construction of an organized facial expression representation space. By integrating non-manual gestures into the sign language message with the assistance of sentiment and semantic data, it ensures both accuracy and richness of the conveyed message. Results of experiments yielded state-of-the-art performance against all existing models on the datasets How2Sign and PHOENIX14T. This research promises much advancement in the area of bridging communication gaps for the deaf and hard-of-hearing community and could inspire further applications in sentiment-aware facial expression synthesis.
Jhuma Sunuwar, Samarjeet Borah, and Aditi Kharga introduce the NSL23 dataset, the first dataset for Nepali Sign Language (NSL)[21], which includes 36 consonants and 13 vowels. The dataset consists of 630 videos, capturing 1205 gestures performed by 14 volunteers, categorized into 9 beginners and 5 experts. Videos were recorded under various environmental conditions—prepared, unprepared, indoor, outdoor, bright, and dark lighting—using Nikon D3200 and iPhone 7 Plus cameras. The dataset will be used to train machine learning models for the recognition of NSL gestures, thus developing automatic NSL translation tools. The dataset is openly accessible on the Zenodo platform and can serve as a significant tool in the enhancement of communication between deaf and non-deaf communities.

This work, titled Empowering Sign Language Communication: Integrating Sentiment and Semantics for Facial Expression Synthesis [22], was proposed by Rafael V. Azevedo, Thiago M. Coutinho, João P. Ferreira, Thiago L. Gomes, and Erickson R. Nascimento. The work is centered around developing a system that improves the production of sign language by generating facial expressions based on sentiment and semantics, a critical need in sign language production to express grammar, emotions, and emphasis. Using a graph convolutional neural network and generative latent optimization, the method produces well-organized facial expression representation. By referring to sentiment and semantic data, non-manual gestures are fused into sign language without an impact on the depth and accuracy of a message. The experimental results showed state-of-the-art performance in comparison to existing models on both How2Sign and PHOENIX14T datasets. These could serve as quite promising advancements at filling communication gaps in the context of the deaf and hard-of-hearing population and potentially become motivating applications within sentiment-aware facial expression synthesis further ahead.

This article is entitled "Audio-synchronized textual enhancement in foreign language pronunciation learning from videos"[23] and is written by Valeria Galimberti, Joan C. Mora, and Roger Gilabert. This study focuses on how L2 captions' textual enhancements in synchronization with the auditory input enhance the pronunciation learning of L1 Spanish/Catalan learners of English. This study evaluates the effect of audio-synchronized captions on learners' attention to phonolexical representations using eye-tracking and lexical decision tasks. Results show that synchronized enhancements significantly enhance audio-visual synchrony, guiding learners' attention to auditory forms. Importantly, learners who received synchronized captions rejected mispronunciations faster, suggesting improved phonolexical accuracy. This study showcases the integration of multimodal input and opens doors to advancing pronunciation for second language learners. This may open a wide door to researching the strategies to synchronize this into an effective process for learning language.

R. Gupta and A. Kumar in the article titled "Indian Sign Language Recognition Using Wearable Sensors and Multi-Label Classification" managed ISL recognition by innovative wearable sensors and multi-label classification.[24]. The work was based on the sEMG and inertial measurement units placed on both the forearms of 10 participants that were recorded during gestures of 100 commonly used ISL signs. Unlike the error-propagating traditional hierarchical tree-based classification methods, the proposed MLC model classifies signs using lexical attributes, such as hand motion symmetry, static or dynamic hand states, and dominant hand use. The work brings in new features to identify one- or two-handed signs, motion symmetry, and static vs. dynamic gestures. Results show that the MLC approach reduces the classification error down to 2.73% compared to the tree-based method, which gave 6.22%. The study "Design and Construction of a Cost-Effective, Portable Sign Language to Speech Translator"[25] by Muhammed Rashaad Cassim et al. introduces an innovative system aimed at making the delivery of communication for the deaf and hard-of-hearing community easier through the conversion of sign language gestures into audible speech.

The system is focused on American Sign Language (ASL) fingerspelling, glove-based approach, integrated with flex sensors, accelerometer, and machine learning algorithms for gesture detection and classification. It achieves an impressive practical accuracy of 85.51% at a cost under $35, much lower than existing alternatives. The solution includes three subsystems: gesture detection, gesture classification, and text-to-speech conversion. Support vector machine with a radial basis kernel is used to enhance the gesture classification. The text-to-speech functionality employed makes use of the eSpeak engine by translating classified gestures into audible output. This research contributes an invaluable piece toward bridging the communication gap between users of sign language and the hearing community, mainly due to portability and cost-effectiveness and accessibility.

## 3 Proposed Methodology

The user input in the proposed audio-to-sign language conversion system with Natural Language Processing could happen at any stage from receiving the user input up to visual content generation in the form of sign language videos or gifs.

The system first receives user input that could be in the form of audio or text. This could be captured through a web interface which is developed with HTML and CSS, and incorporated into a Django-based backend. The first step would be to identify the type of input.
Audio Processing (if audio input)

The system used PyAudio for recording the audio in the required language. The recorded sound is stored in the format of a .wav file. In case the input is audio, this system uses the Google speech recognition api to convert the spoken words (.wav file) to text. This step is critical for handling non-text input because the system will be able to process the audio in any language. Through the conversion to speech to text, the text can be processed by the NLP elements of the system.

| **Title** | **Technology** | **Advantages** | **Disadvantages** | **Performance** |
|---|---|---|---|---|
| Establishing a multimodal dataset for Arabic Sign Language (ArSL) production Samah Abbas et al.[3] | Dataset creation for Arabic Sign Language (ArSL) production and evaluation | Multimodal dataset creation, High-quality data for training models | Limited to Arabic Sign Language, Needs large dataset for better performance | Provides a comprehensive dataset for ArSL, Useful for training recognition systems |
| Arabic Sign Language Recognition and Generating Arabic Speech Using Convolutional Neural Network MM Kamruzzaman et al.[16] | CNN-based recognition and speech generation from Arabic Sign Language | Accurate recognition of signs, Converts sign language to speech, High recognition rate | Requires large training data, Complex model for real-time use | High recognition accuracy, Generates speech from signs, Effective for communication |
| VoiceMitra - A Speech to Visual Translation Approach for Disabled Samruddhi Pate et al.[4] | Speech to text, Language Translator, and Sign Language Converter using NLP | Bridges communication gap, Helps disabled individuals interact, Real-time translation | Requires good speech input quality, May struggle with noisy environments | Real-time translation, Effective communication for disabled users |

| | | | | |
|---|---|---|---|---|
| Audio to Sign Language Translator Anju Yadav et al.[9] | Speech recognition, NLP, ISL dictionary for animation generation | Real-time translation, User-friendly interface, Supports audio input for accessibility | Limited to Indian Sign Language, Animation generation may not cover all words | Real-time audio to sign language translation, Improves accessibility for deaf individuals |
| Voice and Speech Recognition in Tamil Language R Kiran et al.[12] | Tamil speech recognition system for mobile applications, Speech-to-text conversion | Enables Tamil speakers to use voice commands, Flexible and user-friendly for native speakers | Limited to Tamil language, May struggle with noisy environments, Requires large training data for accuracy | Effective for Tamil speech recognition, Facilitates voice commands and SMS sending in Tamil. |
| Bi-directional translation methods of ASL and speech/text Amira Elnashar et al.[1] | Camera-based Gesture Recognition,Data Gloves to capture sensor hand movements | Offers high accuracy for predefined gestures,Simplifies complex translations into ASL grammarExpensive and limited in recognizing dynamic gestures. | Expensive and limited in recognizing dynamic gestures,May not fully capture nuances of sign language | High precision for static gestures, Efficient grammar conversion |

**Table 1.** Comparative study table

## 3.1 Text Processing (Employing NLTK)

Once the input is trapped from any form and transcribed into texts, NLTK handles this query; such is its professed power, a Python module that renders it able to conduct natural language processing with ease. Various core NLP approaches work on them to eventually upgrade the textual quality into something meaningful for the functioning of the mapping system using sign language.

**Tokenization**: The words or tokens in text form are separated. This technique reduces the text to its most basic structure.
Stopword Removal: Words with basic meanings, such as "the," "a," "in," exhibit redundancy in influence and are cut out of the text.
**Clipping**: The words are cut down to their stem or basis form; this means that the various word forms are normalized away from the input sequence, so the processing is made that much easier.
**Porter Stemmer:** This method involves removing derivational affixes to arrive at the base form of the word. For example, the word "ran" may have a stem with "run."
**Lemmatization:** Lemmatization is the award of stripping a word off to its root, or base form would be really useful-in this case, "Better" would be considered to be "Word," which enhances the capabilities of text analysis.

## 3.2 Word Alignment and Sign Language Translation

After text preprocessing, the system will perform word alignment of every token or processed word in the built database of sign language. A database that contains video clips or gifs of sign language gestures of frequent words does this. If there's a match for words in input to those in the database, the corresponding sign language video or gif is retrieved and presented to the user.

Direct Match: System displays corresponding sign language animation for all matching words from the input in the database.
No Exact Match: The system only uses a Punkt-influenced Sentence Tokenizer technique to break the word down into individual letters if it doesn't find the term on its database. So, if a word such as "elephant" doesn't exist, it will break this down into "e", "l", "e", "p", "h", "a", "n", "t". Then sign language representations are fetched for every one of these letters.

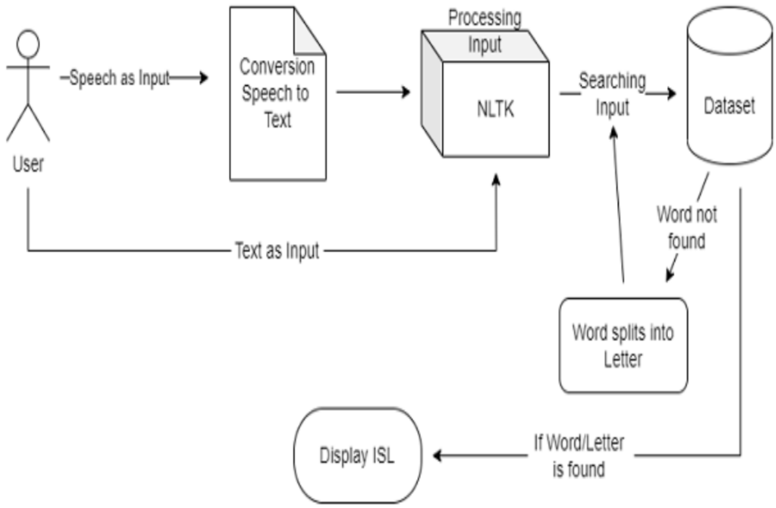


**Fig. 1.** Speak to sign translation methodology

### 3.3 Continuous updating of the database through machine learning integration

The system will continuously update and learn through machine learning. The more data or words, phrases, new sign language videos, etc., it processes, the more words it will recognize and enhance its database. Accordingly, the system will get better with time, proficiently being able to distinguish between languages, accents, and dialects, thus increasing accuracy and coverage.

### 3.4 Security and User Authentication

The system operates safety while maintaining the privacy of the users by implementing authentication mechanisms in order to allow access to the platform. This can be achieved through OAuth or JWT (JSON Web Tokens), which ensures safe login as well as management of a session for the user. Their activities are also logged in so that accountability may be monitored to make the user's experience with the system better.

## 4 Conclusion

The Audio-to-Sign Language Translators are some really keen efforts oriented towards solving becoming lingua franca for hearing impairment society in Indian conditions. Spoken language is thus able to be translated into Indian Sign Language (ISL) gestures manifested through representation in the images or GIF format by a real-time speech recognition and NLP approach. Modular and scalable architecture provide provision for tailoring to different linguistic and cultural contexts, surpassing current solution limits like a restricted vocabulary, slow processing, and insufficient cultural relevance. This system is meant to influence the communication, education, and social participation of hearing-impaired persons. Key features such as real-time processing, high accuracy, and scalability provide opportunities for future advancement. Future work includes vocabulary extension of the existing system, addition of other regional languages, and dynamic generation of gestures using machine learning. Besides, the system will become increasingly usable in multilingual and multimodal contexts, thus amplifying its applicability. This project permutation of novel assistive communication technology by an inclusive design is amplified by cutting-edge technologies. This would pave the way into a more inclusive society while opening up future research in the speech-to-sign translation and human-computer interaction fields. All these prove that, in reality, the information in this paper is based on any technology available up until October 2023 and has not, in fact, been produced anywhere since that time.

## References

1. Amira Elnashar , Karim Hamdan , Sultan Al Seiari , Yusuf Shanableh , Gerassimos Barlas,"Bi-directional translation methods of ASL and speech/text" 2024 ScienceDirect
2. .Biplov Paneru , Bishwash Paneru , Khem Narayan Poudyal ".Advancing human-computer interaction: AI-driven translation of American Sign Language to Nepali using convolutional neural networks and text-to-speech conversion application"2024 ScienceDirect
3. Samah Abbas , Dimah Alahmadi , Hassanin Al-Barhamtoshy "Establishing a multimodal dataset for Arabic Sign Language (ArSL) production Author links open overlay panel" 2024 ScienceDirect

4. VoiceMitra - A Speech to Visual Translation Approach for Disabled, 2023 International Conference on Advanced Computing Technologies and Applications (ICACTA).
5. VijayaJ. VijayaChahat MittalChinmay SinghM.A. Lekhana".An Efficient System for Audio-Based Sign Language Translator Through MFCC Feature Extraction"2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)
6. .Vaddhiraju Swathi, Chada Gautham Reddy , Nomula Pravallika , P. Varsha Reddy , Ch. Prakash Reddy "Audio to sign language converter using Python" 2023 Journal of Survey in Fisheries Sciences
7. Jashwanth Peguda; V Sai Sriharsha Santosh; Y Vijayalata; Ashlin Deepa R N; Vaddi Mounish"Speech to Sign Language Translation for Indian Languages" **2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)**
8. Dr.S.Kumaravel,Mr.G.Santhanakrishnan ""Voice to Sign Language - A Review." 2020
9. .Anju Yadav; Rahul Saxena; Bhavna Saini; Vivek K Verma; Vibhav Srivastava "Audio to Sign Language Translator Web Application" 2021 International Conference on Computational Performance Evaluation (ComPE)
10. R. San-Segundo a, R. Barra a, R. Córdoba a, L.F. D'Haro a, F. Fernández a, J. Ferreiros a, J.M. Lucas a, J. Macías-Guarasa b, J.M. Montero a, J.M. Pardo a " Speech to sign language translation system for Spanish" 2008 ScienceDirectI.A. Adeyanju , O.O. Bello , M.A. Adegboye "Machine learning methods for sign language recognition: A critical review and analysis" 2021 ScienceDirect
11. I.A. Adeyanju , O.O. Bello , M.A. Adegboye "Machine learning methods for sign language recognition: A critical review and analysis" 2021 ScienceDirect
12. Kiran R.; Nivedha K.; Pavithra Devi S.; Subha T. "Voice and speech recognition in Tamil language"2017 2nd International Conference on Computing and Communications Technologies (ICCCT)
13. Mourad Brour a, Abderrahim Benabbou b "ATLASLang MTS 1: Arabic Text Language into Arabic Sign Language Machine Translation System,Second International Conference on Intelligent Computing in Data Sciences (ICDS 2018)
14. Sharada C. Sajjan; Vijaya C"Continuous Speech Recognition of Kannada language using triphone modeling"2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)
15. Dheenesh Pubadi; Ayush Basandrai; Ahmed Mashat; Zvikomborero Chiurunga; Ishan Gandhi; William Ofei "A focus on codemixing and

codeswitching in Tamil speech to text ”2020 8th International Conference in Software Engineering Research and Innovation (CONISOFT)

16..**MM Kamruzzaman** Al-Jouf University “Arabic Sign Language Recognition and Generating Arabic Speech Using Convolutional Neural Network” 2020 Research Gate

17.A. Harkude, S. Namade, S. Patil, and A. Morey, "Audio to Sign Language Translation for Deaf People," *International Journal of Engineering and Innovative Technology (IJEIT)*,

18.H. Mishra, M. Sharma, M. A. Qureshi, and S. Chaudhary, "Audio to Indian Sign Language Translation," Department of Computer Science and Engineering, Krishna Engineering College, Ghaziabad.

19..D. Bhagat, H. Bharambe, J. Joshi, and S. Gul, "Speech to Indian Sign Language Translator," Department of Computer Engineering, Theem College of Engineering, Boisar (E), Mumbai, Maharashtra, India.

20.R. V. Azevedo, T. M. Coutinho, J. P. Ferreira, T. L. Gomes, and E. R. Nascimento, "Empowering Sign Language Communication: Integrating Sentiment and Semantics for Facial Expression Synthesis," Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, and Universidade Federal de Viçosa, Viçosa, Brazil.

21.J. Sunuwar, S. Borah, and A. Kharga, "NSL23 Dataset for Alphabets of Nepali Sign Language," Sikkim Manipal Institute of Technology, Sikkim Manipal University, Sikkim, India.

22.X. Li, S. Yang, and H. Guo, "Application of Virtual Human Sign Language Translation Based on Speech Recognition," School of Computer Science and Engineering, Tianjin University of Technology, Xiqing District, Tianjin, China.

23.V. Galimberti, J. C. Mora, and R. Gilabert, "Audio-Synchronized Textual Enhancement in Foreign Language Pronunciation Learning from Videos," Department of Modern Languages and Literatures and of English Studies, Universitat de Barcelona, 585 Gran Via de Les Corts Catalanes, 08007, Barcelona, Spain.

24.R. Gupta and A. Kumar, "Indian Sign Language Recognition Using Wearable Sensors and Multi-Label Classification," Electronics and Communication Engineering Department, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Sector 125, UP, India, and Centre for Applied Research in Electronics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India.

25.M. R. Cassim, J. Parry, A. Pantanowitz, and D. M. Rubin, "Design and Construction of a Cost-Effective, Portable Sign Language to Speech Translator," Biomedical Engineering Research Group, School of Electrical and

Information Engineering, University of the Witwatersrand, Johannesburg, Private Bag 3, 2050, Johannesburg, South Africa.