Audio to Sign Language Translator Web Application

Anju Yadav
Dept. of IT
(Manipal University Jaipur)
Jaipur, India
anju.yadav@jaipur.manipal.edu

Rahul Saxena
Dept. of IT
(Manipal University Jaipur)
Jaipur, India
rahul.saxena@jaipur.manipal.edu

Bhavna Saini
Dept. of IT
(Manipal University Jaipur)
Jaipur, India
bhavna.saini@jaipur.manipal.edu

Vivek K Verma
Dept. of IT
(Manipal University Jaipur)
Jaipur, India
vivek.verma@jaipur.manipal.edu

Vibhav Srivastava
Dept. of IT
(Manipal University Jaipur)
Jaipur, India

Abstract— Sign language is an effective mode of conversation for persons who have difficulty speaking or hearing. There are numerous media accessible for translation or for identifying sign languages and converting those to text format, However, methods for converting text to sign language have been few and even not web-based software, owing to the scarcity of resources. The proposed web application seeks to develop a translating mechanism or automation that includes a parser element that converts the incoming speech data or English text to a phrase structure grammar representation, which is then used by another module that contains Indi Sign language grammatical format. This is accomplished through the means of removing stop-words from the reordered input format. Because Indian sign language does not provide word inflections, stemming and lemmatization are used to turn words into their root form. Following sentence filtration, all words are tested against the words in the database, which is represented as a dictionary comprising video representations of each word. If the words are missing from the database, the algorithm will then look for its related synonym and replace it with that term.

In many ways, the proposed system is more innovative and efficient than existing systems, because Existing methods can only convert words directly into Indi sign language, and they were not as efficient as this system, whereas this in the actual world, the system tries to translate these phrases into Indian sign language grammatical order. Because this is a web-based programmed, it is straightforward to access and use. This technology is platform agnostic and more versatile to use, and it transforms phrases to sign language in real time.

Keywords— Communication Interpreter, Sign Language, Deaf-Mute, Web Application, Audio to Sign Language, ISL

I. INTRODUCTION

In this paper Natural language processing (NLP) is used to transform the English text/words language or voice input to Indian Sign Language (ISL) and integrate it in a web-based software to improve the communication capabilities of people with hearing impairment. Sign Language (SL) is a common technique that conveys meaning through hand/manual gestures, hand gestures, and other modalities such as facial expressions and physical movements. In this paper we have employed videos or a database that has been particularly developed for specific words to translate text or voice input into SL [15].

According to the estimation, 7105 known living language all around the world, which are organized into 136 language families. One of these 136 families is SL, which is currently in use by people with impaired hearing to transmit their message. This language includes around 136 SL from over the globe, depending on area. Approximately 72 million of the world's nearly 7 billion inhabitants are deaf. Around 4.3M people utilize SL out of such a big number. The remaining over 67M absolute deaf and partially deaf people do not use SL to communicate. As a result, over 90% of deaf people have very little or no access to education and other information [1][2].

A word in spoken language is made up of phonemes. Two words are differentiated by at least one phoneme. A sign in SL is made up of cheremes, and two signs can differentiated with at least one chereme [4]. A simultaneous or sequential creation of its manual and non-manual cheremes is an action/sign. Several criteria can be used to define a manual chereme, including hand form, hand placement, hand motion/orientation, and limb movements (straight, circular or curved). Facial expressions, eye focus, and head/body posture are some of the elements that identify non-manual chereme [3][8]. ISL hierarchy is discussed in Fig. 1
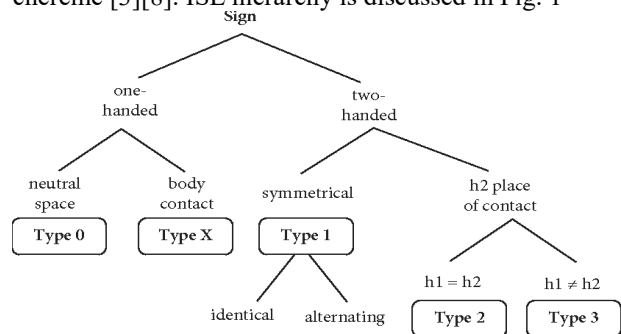

Fig. 1. Indian Sign Language Type Hierarchy

One Handed Signs: A single dominant hand represents the one-handed signs. A stationary/static or dynamic one-hand symbol is possible (having movements). Each static and moving sign is further divided into two categories: manual and anti-manual. The image shows, for instance: one-hand static signage with anti-manual and manual compositions.

Two Handed Signs: The signer's hands are used to symbolize the both-hand/two-hand signs. The classification of double hand signs is identical to that of single hand signs. Both hand signals with motion, on the other hand, extends as Type0 and Type1 signs [5][6].

Type0 signs are those where both the hands are engaged.

Type1 As illustrated below, gestures are those where the one hand (dominant) is even more active than that of the other see Fig. 2.

Fig 2. ISL Hand gestures

It's crucial to clarify that SL aren't "natural languages represented by signs" or even "word for word" converted into signs. The term light, for example, has numerous meanings in English. Light means "not too heavy," or "light color," which means "not too dark" or "switch on the light". In English, we are using the same word 'light' to convey these different meanings, but in SL we use different terms. As a result, SL depicts the meanings of words rather than the words themselves [9].

Any language in which each word is represented by a sign is known as sign lang. English word order is used only in Signed English, and each word is represented by a sign. Sign A sign is not displayed for every word in Supported English; instead, signs are displayed for select of the most important terms. A further version is Hand Phrasing, which uses fingers to display the letters that makes up the word instead of symbols. This type is mixed with any other form of SL due to the restricted SL lexicon [10].

II. LITERATURE REVIEW

In literature many authors have worked on American Sign Language (ASL), ISL, and other SL, according to the literature. In comparison to ISL, the amount of ASL research available on the internet is vast. The major reason that ISL is not widely used in research is a lack of dictionaries and exposure.

The detailed study on ISL is discussed in this section, however the mythology is different. In these research papers, sigmal languages and software-based translation are used, which are inconvenient and difficult to use. Furthermore, the approaches are not highly optimized or user pleasant. The methodology in proposed work is user-friendly and web-based, so it can be accessed from any device and is quick and dependable.

Khushdeep et.al. [19]	Using sigmal notation for sign language representation.
Goyal et al. [16]	Synthetic animation is used for conversion.
Praveen et.al. [20]	For each every word, a video will be stored in a preset database.
Purushottam et al. [22]	Generation of Sign language with signing Avatar.
Lalit & Vishal [7]	SiGML (Signing Gesture Markup Language) are used as XML encoding of HamNoSys.)
Madhubala et al. [21]	Sign language and the text will be converted to sign images.
Martin et al. [14]	Creation of corpus and generating an avatar for disaster domain are also applicable to other domains as well
Shinde et al. [23]	The dataset represents a significant variety of hand motion pictures gathered from internet.
Kishore [18]	Fuzzy logic for recognition of gestures.
Anuja [17]	ISL from images or video sequences.

ISL is a frequently used communication aid for those who are deaf or hard of hearing. There are no efficient software programmers that can translate text to ISL. India has had relatively little work done on computerization and development of SL in recent years [16].

So far, software throughout this field has been mostly concentrated on either ASL or British SL (BSL). Translation, while systems or software produced for ISL translation are scarce and ineffective. Most of the systems' fundamental architecture is based on:

Direct translation: Terms in the source language are immediately translated into target language words. It's possible that the result won't meet your expectations.
Static. Machine Transformation: It needs a large parallel database, which is not easily available in the case of sign language.
Transfer based architecture: To specify a correct translation from one language system to another, grammar rules are used.

III. FACTS ABOUT ISL

SL is a natural language that contains several facts that most people are not having much knowledge of these. The following MT systems are used for other SL see Fig. 3:
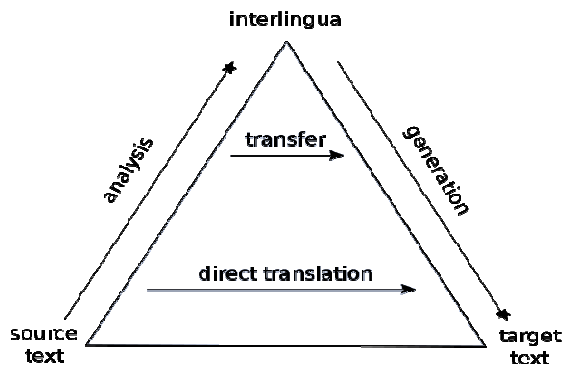
Fig. 3. MT systems used for other sign language

- **Direct conversion/translation:** It essentially turns words into SL. The grammar is unaffected, no syntactical operation is performed on the original text, and so the order of conversion is untouched.
- **Transfer Based:** In these systems, the source language is turned into some intermediate output, which is subsequently translated into the target language using some linguistic rules. It was also termed as rule-based conversion.
- **Interlingua translation:** In this computational process is then used to construct the target language. As a result, it provides an option to both direct and transfer-based translation [11].

IV. VISIONS TOWARD THE PROBLEM

This system aims at conversion; it will assist hearing-impaired people in understanding in an effective and straightforward manner by presenting them with a video animation to transmit the message of text or voice. This project intends to cover the domain of ISL, which has around 1500 words in its lexicon. Each of these terms will have a video associated with it. Words that do not appear in the dictionary would be substituted with synonyms based on the delicacy of the words as well as their parts of speech [12][16].

A. Proposed methodology

A few studies been conducted to develop some model that is based on the above ideals described above and caters to Indian sign language. As a result, it was recommended to create one for ISL using transfer-based translation and to deploy it as a website. This translation model's performance and accuracy will be totally dependent on the conversion of voice input to English. Text and then to ISL following the algorithm given.

This approach is distinct since the proposed methodology use the concept of video database, whereas other strategies employ images or animations to communicate the content. Proposed work also includes the creation of modules containing the logic of stop words and other computations on sentences. Stop-words have been deleted, reducing text processing efforts.

B. Hypothesis

A bilingual dictionary is required for translation from source to target language. As a result, we will create such type of

dictionary that includes both the Eng. term along with comparable Indian sign. Images, films, or coded sign language text could be used to represent the English word's counterpart (gloss). All of the methods have its advantages and limitations; however, the video method is the best because it outperforms another algorithms. From the official website, we will save visual content of every syllable. Further, evaluated every clip independently and delete others who we deem inappropriate. Whereas will retain some raw inputs with a diverse set of syllables [7].

C. Algorithm design

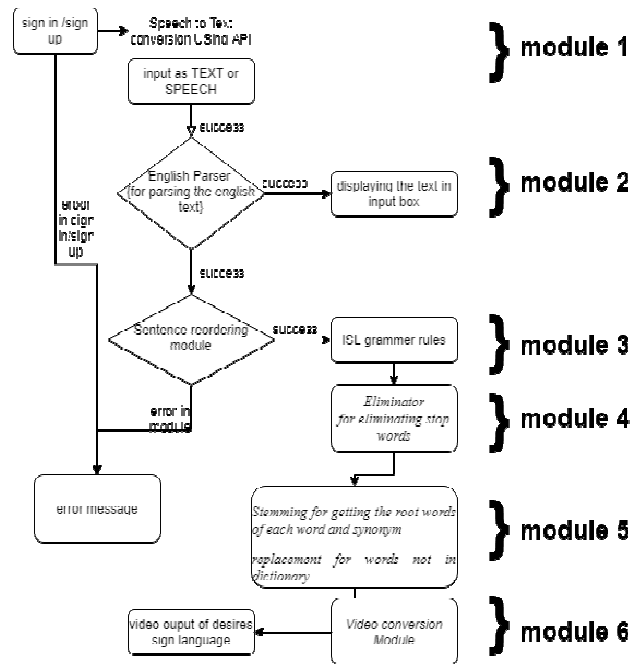The system is designed with six modules for the algorithm see Fig. 4:

Fig. 4. Algorithm of the ISL web application

The algorithm receives typed English text, which is then reordered to match needs of ISL grammar. The module of the English text is reordered so that undesirable words are deleted. When an English word corresponds to an ISL It can be perceived as a real person video, a sign image, or a simulation.

In this translation system, video messages have indeed been generated for each appropriate Word in English. A video is obtained from the internet and translated into animated action using the blender tool to make animations on each English word. We used all the ISL grammar and constructed a database of all conceivable terms with hard gestures and facial gestures see Fig. 5.
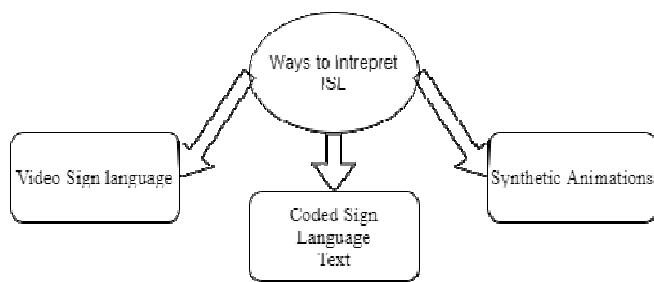
Fig. 5. ISL interpreter

Parsing of the input English sentence

Module 1: Preprocessing is a procedure for measuring a sentence's grammatical structure. To extract the conjugation of English sentences, 3rd party application is used, which utilizes syntactical grammatically modeling, as in the directive method. To translate these into the sentence per the language form of the English word, the grammatical constructions of the English word was essential.

Grammar rules for conversion of English sentence to ISL sentence

Module 2: It converts an English sentence to an ISL sentence utilizing 20 distinct verb patterns in the data. All the verb patterns (20 patterns) are examined to transform an English sentence to an ISL sentence.

Elimination of unwanted words

Module 3: Indian Sign Language (ISL) sentences are made up of key words, from which any unnecessary words must be removed. These words include TO, POS, MD, FW, CC, JJS, NNS, NPS, SYM, RP and SYM. These grammatical forms are not used in ISL translation and hence the module designed is to eliminate all those words which cannot be used for ISL conversion [7].

Lemmatization & synonym replacement

Module 4 and 5: ISL is a nonverbal communication that is commonly used in India. There must be no suffix, auxiliary verbs, or inflexions in any of the words given. Whenever a word in an ISL sentence is not a root word, it's lemmatized by the recognition system see Fig. 6 [13].

Fig. 6. Stemming and Lemmatization

Sign video representation

Module 6: ISL speech signal after finishing the prior processes, and the algorithm will search the database for similarities with each of the phrases. It will be predicated on a basic string-matching approach that will be used to both

processed text corpus and video classification. Finally, a representation of a succession of clips can be seen on the screen, one after the other.

The result of this module will be a short video with ISL interpreted phrases. Each phrase will have its own video in the library, and the final output will be a compilation of these videos.

V. RESULT

When the database is first started, the dictionary, which contains all the videos, is uploaded to the server, and the system starts up smoothly. This programme is tested against a variety of specified test cases, and it is almost ready to go online.

It is tried for both basic and complicated test cases, including birder test scenarios such as all the words with no meaning, and the programmed responded effectively by delivering word-by-word output to the challenge. Because sign language does not implement tense, we mandate the framework to lemmatize each of the words. However, lemmatization obeys a dataset, and we chose the Wordnet repository, which does not essentially stem each of the verb forms, necessitating the modification of video and display.

Instead of displaying separate words using a GIF/ image, our seamless integration transforms the whole input text into a single visual, giving the model a far more realistic and dynamic appeal. Because the ISL database is still limited and continues to expand, much more work may be accomplished on this track.

Test cases result compilation:
- Finding relevant films and putting them together in a collaborative is a time-consuming process.
- Due to a lack of study, the video output could not guarantee uniformity on this field
- Because all words do not always have an equivalent word file, such terms are skipped.
- Video playback is seamless and quick.
- Videos have a realistic feel but still lacks a proper movement.

Border line test cases:
Oh! and other exclamation words are examples of exclamation phrases. Alas! Alternatively, yay! Although there is no accurate translation in ISL, those were words that are focused on visual expressions and gestures and must be handled appropriately.

Imperative expressions, including such "book the flight," can be transformed onto ISL grammatical style and stated as "Flight Book," but they cannot be stated in the imperative format due to the defined structure.

The website's visual representation looks like this, with animations on the right side and all input and visual data on the left. Videos will be played one after the other without any pauses or stops, with the option for the user to pause or replay at any time.

According to the hypothesis, the software's overall performance is good, and it passed 80 percent of the test cases. The programme is small and easy to use, and it may be used on any device. In comparison to earlier versions of software, the video representation is smooth, and the system is more resilient.

Fig. 7. Reading the voice input and text input

Fig. 8. Processing of the speech

Fig. 9. Video of the speech processed from voice and text

This is the software's real representation; there is a microphone button in Fig. 7 that activates the user's *voice input*, as well as a *text input* bar where the user may enter text. Following successful text entry, a submit button appears, allowing the programme to do the necessary actions on the text and display the finished text to be processed for video conversion. The text we enter by speaking or typing is processed and presented in the module in this Fig. 8. The text is presented in two columns: the actual text column and the text after all actions.

As soon as the user clicks the play/pause button, the video that is queued for show appears on the right side of the software and runs without any more input, displaying the needed message see Fig. 9.

By pre-processing the text information using Natural language techniques and utilizing the necessary phrase as from feature vector, low weight clips enhance system results. Additionally, unlike many past initiatives, we have employed graphic portrayal for the whole input sentence rather than just a single word.

VI. CONCLUSION

This paper presents a technique for translating English text into Indian sign language. The aspects is truly revolutionary for Indian sign translation software in the modern world. The transformation unit (that utilizes grammatical structure to convert Correct language to ISL sentences) and the exclusion module are the two primary elements of effective system (which removes undesirable words from ISL sentences). There will be video converter components, as well as substitution and lexicon - based interfaces (that translate all phrase during an ISL expression to its root word meaning) (converts the ISL sentence to video film representation). The system is actively being designed to interpret English translation into synthesized animations containing video representation. The thesaurus of Words and phrases to Indian signs, wherein the pseudo component of each sign is supposed to take care of, including the character's mouth being rendered so a hearing-impaired individual could interpret the lips, requires a considerable amount of time and commitment. Numerous impaired schools were used to validate the software's overall conversion efficiency.

REFERENCES

1. Thompson, S. (1993). WFD Survey of Deaf People in the Developing World.

2. Zeshan, U., Vasishta, M. N., & Sethna, M. (2005). Implementation of Indian Sign Language in educational settings. Asia Pacific Disability Rehabilitation Journal, 16(1), 16- 40.

3. Zeshan, U. (2003). Indo-Pakistani Sign Language grammar: a typological outline. Sign Language Studies, 3(2), 157- 212.

4. Stokoe, W. C. (2005). Sign language structure: An outline of the visual communication systems of the American deaf. Journal of deaf studies and deaf education, 10(1), 3-37.

5. Klein, D., & Manning, C. D. (2003, July). Accurate unlexicalized parsing. In Proceedings of the 41st Annual Meeting on Association for Computational Linguistics Volume 1, 423-430.

6. Hanke, T. (2004, May). HamNoSys-representing sign language data in language resources and language processing contexts. In LREC , Vol. 4, 1-6.

7. Goyal, L., & Goyal, V. (2016). Development of Indian sign language dictionary using synthetic animations. Indian journal of Science and Technology, 9(32), 1-5.

8. Stokoe, W. C. (1978). Sign language structure: The first linguistic analysis of American Sign Language. Linstok Press.

9. Sutton, V. (2014). Lessons in sign writing: Textbook. SignWriting.

10. Sinha, S. (2003). A skeletal grammar of Indian sign language. Unpublished master's diss., Jawaharlal Nehru University, New Delhi, India.

11. Aboh, E., Pfau, R., & Zeshan, U. (2005). When a wh-word is not a wh-word: The case of Indian Sign Language. The yearbook of South Asian languages and linguistics, 2005, 11-43.

12. Wren, P. C., & Martin, W. (2005). High school English grammar and composition. S Chand.

13. Porter, M. F. (2001). Snowball: A language for stemming algorithms. Disaster Message Using Avatar

14. Martin, P. M., Belhe, S., Mudliar, S., Kulkarni, M., & Sahasrabudhe, S. (2013). An Indian Sign Language (ISL) corpus of the domain disaster message using Avatar. In Proceedings of the third international symposium in sign language translations and technology (SLTAT-2013) (pp. 1-4).

15. Shinde, A., & Kagalkar, R. (2015). Sign Language to Text and Vice Versa Recognition using Computer Vision in Marathi. International Journal of Computer Applications, 975, 8887.

16. Goyal, L., & Goyal, V. (2016, December). Automatic translation of English text to Indian sign language synthetic animations. In Proceedings of the 13th International Conference on Natural Language Processing (pp. 144-153).

17. Nair, A. V., & Bindu, V. (2013). A review on Indian sign language recognition. International journal of computer applications, 73(22).

18. Kishore, P. V. V., Kumar, P. R., Kumar, E. K., & Kishore, S. R. C. (2011). Video audio interface for recognizing gestures of indian sign. International Journal of Image Processing (IJIP), 5(4), 479.

19. Kaur, K., & Kumar, P. (2016). HamNoSys to SiGML conversion system for sign language automation. Procedia Computer Science, 89, 794-803.

20. Praveen, N., Karanth, N., & Megha, M. S. (2014, October). Sign language interpreter using a smart glove. In 2014 International Conference on Advances in Electronics Computers and Communications (pp. 1-5). IEEE.

21. Madhubala, D., Rao, N. K., Roopa, D. Y. M., Kalyani, D., Pranay, M., & Raju, C. K. A tool to convert audio/text to sign language using python libraries. Turkish Journal of Physiotherapy and Rehabilitation, 32, 2.

22. Kar, P., Reddy, M., Mukherjee, A., & Raina, A. M. (2007). Ingit: Limited domain formulaic translation from hindi strings to indian sign language. ICON, 52, 53-54.

23. Shinde, A., & Kagalkar, R. (2015). Sign Language to Text and Vice Versa Recognition using Computer Vision in Marathi. International Journal of Computer Applications, 975, 8887.