



Article

# Translating Speech to Indian Sign Language Using Natural Language Processing

Purushottam Sharma <sup>1,\*</sup>, Devesh Tulsian <sup>1</sup>, Chaman Verma <sup>2,\*</sup>, Pratibha Sharma <sup>1</sup> and Nancy Nancy <sup>1</sup>

<sup>1</sup> Amity School of Engineering & Technology, Amity University, Noida 201301, India

<sup>2</sup> Department of Media and Educational Informatics, Faculty of Informatics, Eotvos Lorand University, 1053 Budapest, Hungary

\* Correspondence: puru.mit2002@gmail.com (P.S.); chaman@inf.elte.hu (C.V.)

**Abstract:** Language plays a vital role in the communication of ideas, thoughts, and information to others. Hearing-impaired people also understand our thoughts using a language known as sign language. Every country has a different sign language which is based on their native language. In our research paper, our major focus is on Indian Sign Language, which is mostly used by hearing- and speaking-impaired communities in India. While communicating our thoughts and views with others, one of the most essential factors is listening. What if the other party is not able to hear or grasp what you are talking about? This situation is faced by nearly every hearing-impaired person in our society. This led to the idea of introducing an audio to Indian Sign Language translation system which can erase this gap in communication between hearing-impaired people and society. The system accepts audio and text as input and matches it with the videos present in the database created by the authors. If matched, it shows corresponding sign movements based on the grammar rules of Indian Sign Language as output; if not, it then goes through the processes of tokenization and lemmatization. The heart of the system is natural language processing which equips the system with tokenization, parsing, lemmatization, and part-of-speech tagging.



**Citation:** Sharma, P.; Tulsian, D.; Verma, C.; Sharma, P.; Nancy, N. Translating Speech to Indian Sign Language Using Natural Language Processing. *Future Internet* **2022**, *14*, 253. <https://doi.org/10.3390/fi14090253>

Academic Editors: Filipe Portela and Paolo Bellavista

Received: 18 July 2022

Accepted: 23 August 2022

Published: 25 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Every country has a different sign language which is based on their native language. It is not easy for us to speak when we know the other person is not listening, let alone hearing-impaired. Even we with sufficient hearing abilities tend to ignore or avoid communication with those who do not hear, and for those who cannot hear it becomes even more difficult. Having the skill to talk to those who cannot hear can not only bridge the gap between the two but also help in the exchange of a lot of ideas and new thoughts which could encourage these people to contribute to the development of technology. Every mind can contribute to making unknowns into knowns and impossible possible.

### 1.1. Indian Sign Language

Indian Sign Language can facilitate people to create an inclusive society in which people with disabilities have equal chances for growth and development so that they can live productive, safe, and dignified lives. In India's hard-of-hearing community, Indian Sign Language (ISL) is widely utilized. However, ISL is not utilized to teach hard-of-hearing students in deaf schools. Teacher education programs do not train teachers to use ISL in their classrooms. Sign language is not included in any of the teaching materials. The parents of hard-of-hearing children are often unaware of sign language's value in bridging communication gaps. ISL interpreters are in high demand at institutes and other locations where hard-of-hearing and hearing individuals communicate, yet India only has about 300 licensed interpreters.

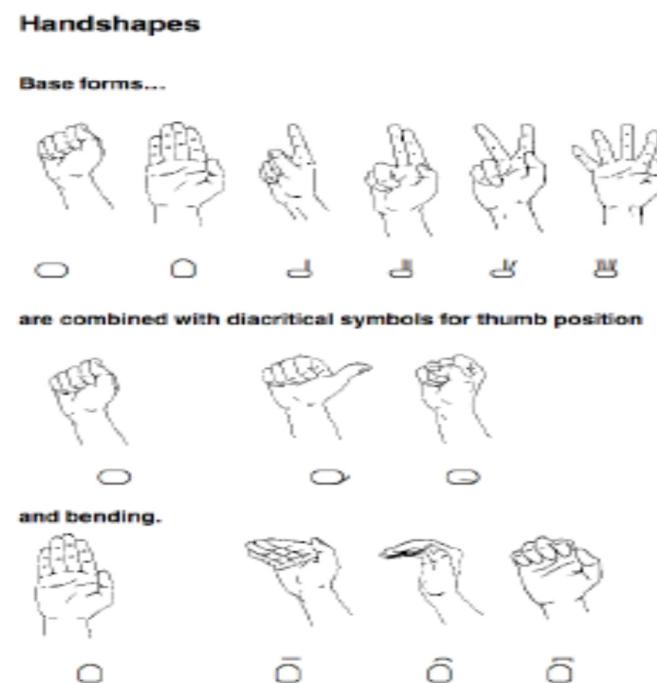
ISL aims to conquer the following points:

- To train people to use Indian Sign Language (ISL) and educate and do research on the language, including bilingualism.
- To encourage hard-of-hearing students in primary, intermediate, and higher education to use Indian Sign Language as a form of instruction.
- To educate and train diverse groups, such as government officials, teachers, professionals, community leaders, and the public, on Indian Sign Language and how to use it.
- To promote and propagate Indian Sign Language in collaboration with hard-of-hearing groups and other institutions working on disabilities.

### 1.2. HamNoSys V/S ISL Gestures

Unlike other sign language scripts, HamNoSys is not intended to be a practical writing tool for everyday communication. Rather, its motivation is similar to that of the International Phonetic Alphabet, which is intended to consistently transcribe the sounds of any spoken language [1]. Because the increased number of possible parameter versions precluded the use of a well-known alphabet, newly produced glyphs must be developed in such a way that memorizing or deducing the meaning of the symbols is as simple as possible [2].

HamNoSys consists of handshapes and gestures as shown in Figure 1.



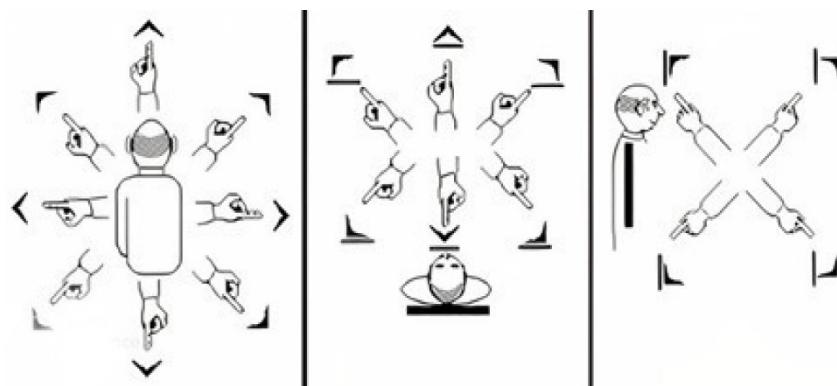
**Figure 1.** Various Handshapes in HamNoSys.

Indian Sign Language consists of gestures as shown in Table 1.

**Table 1.** ISL Handshapes.

One Hand	Open Hand	Zero Hand	Five Hand
U-hand	L-hand	Y-hand	C-hand
V-hand	C-hand	Full U-hand	Full C-hand
Fist hand	Claw hand	Closed two hands	Closed four hands

HamNoSys contains pointers as shown in Figure 2.



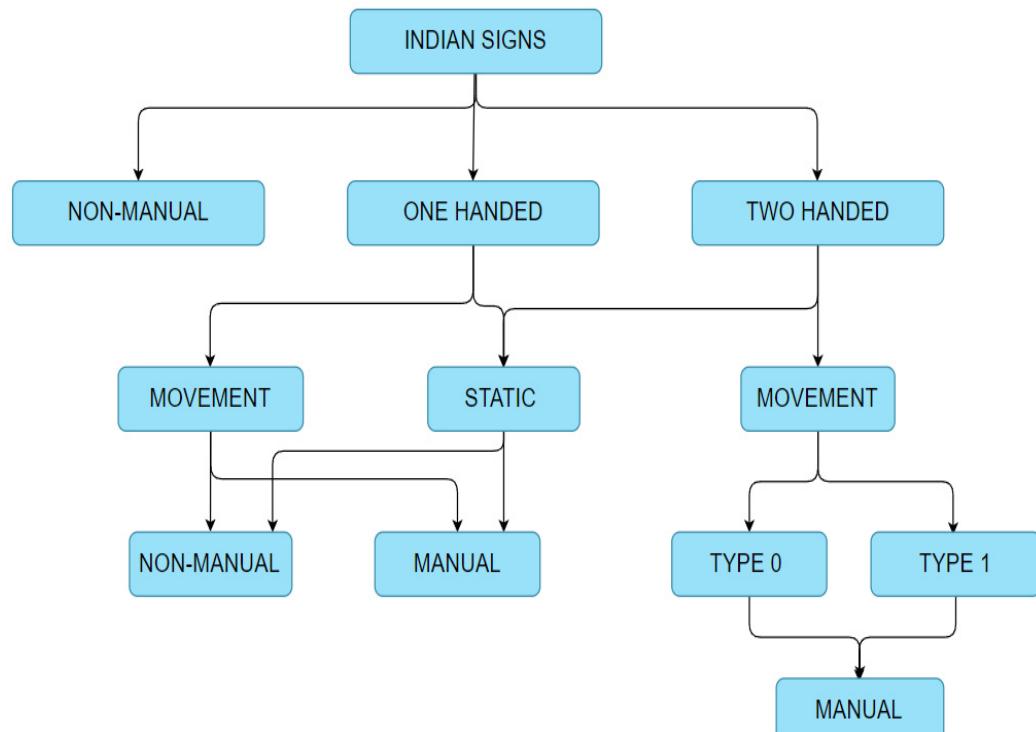
**Figure 2.** Various Directions of HamNoSys Pointers.

Indian Sign Language consists of gestures as shown in Table 2.

**Table 2.** ISL Directions.

Hand Moving UP	Hand Moving DOWN
Moving left	Moving right
Hand moving forward in a semicircle	Hand moving backward in a semicircle
Moving left to right (slant)	Moving right to left (slant)
Moving in a vertical circle	Moving in a horizontal circle
Hand moving UP and DOWN (repeatedly)	Hand moving side by side (repeatedly)

Sign language [3] can be defined as not just hand gestures to express words or sentences but also their meanings. Moreover, when it comes to defining sign language in depth, it can be explained in three broad categories, namely (as described in Figure 3): nonmanual, one-handed, and two-handed.



**Figure 3.** Types of Hand Gestures.

When it comes to body gestures, as the name suggests, it includes body movements other than hand gestures, such as head movements, facial expressions, mouth shapes, etc. and taken as modulation, when speaking we modulate our voice to show the variation in our speech.

Talking about one-handed, it is explained as static and dynamic movements which are further divided into manual and nonmanual.

Two-handed signs are divided into the above sections too, but movements can be further divided into Type 0 and Type 1. Explaining the terms, Type 0 includes the signs which are performed by both hands, whereas Type 1 includes the signs which too involve both hands, but one hand is dominant over the other, taking the lead.

## 2. Literature Survey

### 2.1. Sign Language in English

A two-way communication system is suggested in the paper [4] but the authors are only able to convert 26 alphabets and three characters with an accuracy rate of 99.78% using CNN models. The authors only suggest that future work must be conducted in the field of natural language processing to convert speech into sign language.

In the paper [5], the authors propose a system that converts sign language into English and Malayalam. The authors of the paper suggest using an Arduino Uno, which uses a pair of gloves to recognize the signs and translates the signs from ISL to the preferred language. The system is useful, as it recognizes two-hand and motion signs.

The Indian Sign Language interpreter presented in the paper [6] uses hybrid CNN models to detect multiple sign gestures and then goes on to predict the sentence that the user is trying to gesture by using natural language processing techniques. The system is able to achieve 80–95% accuracy under various conditions.

In another study, the HSR model is used by the authors in converting ISL signs into text. The HSR model gives an advantage over RGB-based models, but this system has an accuracy ranging from 30 to 100% depending upon the illumination, hand position, finger position [7] etc.

The authors of [8] paper propose a system that recognizes 26 ASL signs and converts them into English text. They use principal component analysis to detect the signs in MATLAB.

The ASL to sign language synthesis tool uses VRML avatars and plays them using a BAP player. The major problem with the system is that many complex movements are not possible using the current VRML avatars. For example, touching the hand to any part of the body is not possible in the current system [9].

In another study mentioned in [10], one video-based sign language translation system converts signs from ISL, BSL, and ASL with an overall accuracy of 92.4%. The software utilizes CNN and RNN for the real-time recognition of dynamic signs. The system then converts the signs into text and then uses text–speech API to give an audio output to the user.

The authors of another paper first use the Microsoft Kinect 360 camera to capture the movement of the ISL signs. A unity engine is used to display the Blender 3D animation created by the authors. Although the system can successfully convert words into sign language, it is not able to convert phrases/multiple words into ISL.

### 2.2. Sign Language in Other Languages

The work presented by the authors of [11] is another bidirectional sign language system. The system is able to achieve 97% accuracy when translating sign languages to text or audio. Authors use Google to API to convert speech to text and then the system produces a 3D figure using the unity engine after extracting keywords from the input [12].

Another system proposed by the authors in the paper [13] converts Malayalam text and gives a 3D animated avatar as the sign language output. The system uses HamNoSys notation, as it is the main structure of the signs [14]. A unique Russian text to Russian

sign language [15] system utilizes semantic analysis algorithms to convert text to sign language. It focuses on the lexical meanings of the words. Although the system can reduce the sentence into gestures, it is observed by the authors that the sentence proposition can be improved by further making the algorithm more efficient.

### 3. Comparison

As shown in Table 3, most of the existing models which convert English text to a sign language [16], whether it is BSL, ASL, or ISL, use natural language processing. The major problem which almost all the existing models face is that the conversion of text to sign language only happens if the sign language for that particular word is present in the database.

**Table 3.** Literature Survey.

Source	Objective	Methodology	Conclusion
Sign Languages in English			
[17]	Converting the English language to SIGML representation.	The system uses NLTK for the direct mapping of English text to the HamNoSys string.	The system successfully converts input to a small set of HamNoSys strings and then represents them using SIGML.
[18]	Converting English text to ISL using LOTS notation.	The system utilizes various NLTK processes to carry out the conversion.	The system is successful in showing English text in LOTS notation.
[19]	Converting Indian Sign Language to English text.	The system uses various convolutional neural network techniques.	The system is able to convert sign language to the English language for a very small dataset.
[20]	Converting English to animated sign language as per American Sign Language.	The system takes input by using various speech recognition software and then uses various algorithms to produce animated signs.	The proposed system is still in development and will map each word directly.
[21]	A system to translate English to ISL.	The authors use various NLP and Google APIs.	The system takes input in English and with a 77% accuracy and it converts the input to SIGML animation, but the system is only shown to convert words or letters.
[22]	Generating HamNoSys signs from the user's input.	The authors use the direct mapping of only 100 words to generate the HamNoSys signs.	The system successfully generates SIGML animation for 100 words in the ISL dictionary.
[23]	A two-way communication system between English and ISL.	The authors use neural network techniques and the HMM model.	The authors only describe the English to ISL in this paper and the system only generates output if the input is an exact match with their database.
[24]	WordNet-based English to ISL gloss.	The authors utilize various NLP techniques and the hidden Markov model.	The system removes the words for which it does not find any replacement in the current database. In addition, the system is the only text to text conversion.
[25]	A two-way communication system between hard-of-hearing and hearing people.	The authors utilize various machine learning algorithms and NLP techniques.	The system only shows ISL signs for those words which are stored in the database and skips any other word.
Sign Languages in Other Languages			
[26]	Converting Arabic text to Arabic Sign Language.	The system uses various text parsing and word processing techniques.	The system successfully converts Arabic text to Arabic sign language with 87% efficiency and then shows the corresponding sign language animatedly.

Our model not only overcomes the problem mentioned in the previous paragraph but also takes it one step further by converting and displaying sign language for phrases/sentences, i.e., if the input from the user has any combination of words for which there is a particular sign present in the ISL database then the proposed system will display the sign for those combinations of words in one go.

#### 4. Proposed Work

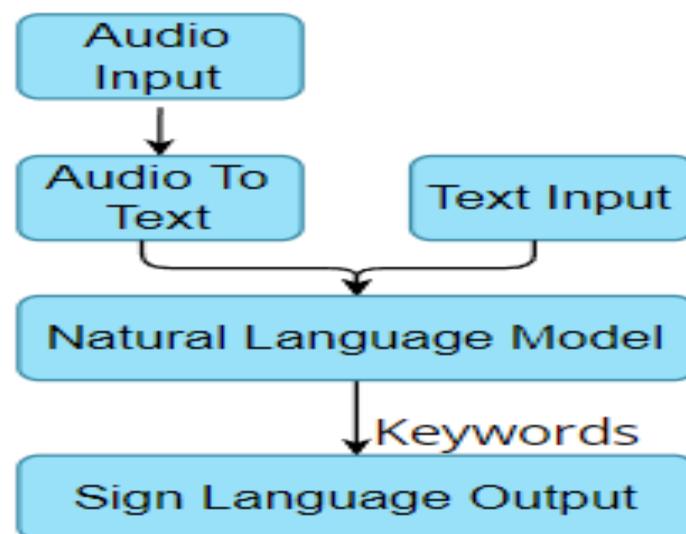
The proposed system presented in this paper is a real-time audio to Indian Sign Language conversion system which will assist hearing-impaired people to communicate easily with other hearing people. The system comprises mainly six components:

1. Audio-to-text conversion if the input is audio.
2. The tokenization of English text into words.
3. Parsing the English text into phrase structure trees.
4. The reordering of sentences based on Indian Sign Language grammar rules.
5. Using lemmatization along with part-of-speech tagging so that synonyms of words or the root form of a word can be used if the exact word is not present in the database.
6. Indian Sign Language video output.

The overall efficiency of the system is improved, as it splits the word into letters. If the video for the corresponding word is not present in the database then it shows video output letter by letter so that it does not skip any word. Another unique feature of the system is that it can recognize phrases in the sentence and show sign language videos corresponding to the phrase if present in a database instead of word by word. The database contains around 1000+ videos which are a combination of self-recorded videos by the author and open-source ISL faculty videos. Thus, it increases the horizon and literacy of the system.

#### 5. System Architecture

Figure 4 shows the system architecture. The user has an option to enter the input either by text or by audio. The input is processed through the natural language model designed by the authors and keywords are given as the output. If the text within the keywords contains phrases or a combination of multiple words for which the sign language video is present in the database, then those videos are shown for such cases otherwise, the keywords are tokenized further into words or letters.



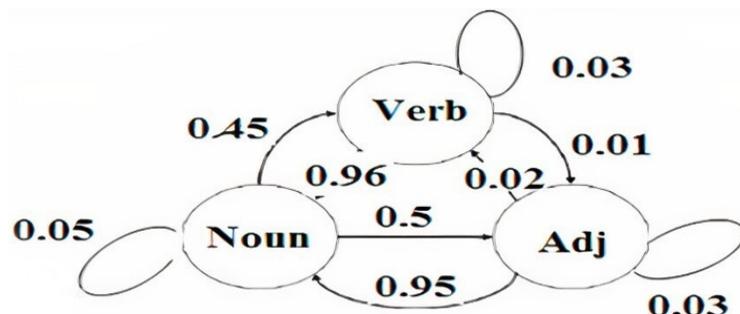
**Figure 4.** System Architecture.

#### 6. Hidden Markov Model

The hidden Markov model is one of the models that may be used as a classifier; it consists of a set of states where the transition from one state to the next is determined by a

specific input. As a result, the shift from state to state continues until the output state or observation is reached. Furthermore, the likelihood of a specific transition is influenced by the likelihood of a transition from before a current condition. The probability model is made up of three basic components: well-defined experiments, a sample space ( $\Omega$ ) that contains all possible events, and an event that is chosen from the sample space. HMM, on the other hand, is based on conditional probability, which implies that the likelihood of a specific event X occurring relies on the probability of a previous event Y occurring [27]. This conditional probability can be expressed as follows:  $P(X|Y) = P(X \cap Y)/P(Y)$ .  $p(X|Y) = P(X \cap Y)/P(Y)$ .

Assume we have three states: verb, noun, and adjective, as indicated in Figure 5. Table 4 shows the transition matrix that shows the transition from one state to another. The transition probability from verb to adj is 0.01 and from adj to verb is 0.02.



**Figure 5.** Markov Model with three states.

**Table 4.** Transition Matrix (Probability of Transition).

State	Next State		
	Verb	Noun	Adj
Previous (current) state	Verb	$P(\text{verb}   \text{verb}) = 0.3$	$P(\text{Noun}   \text{verb}) = 0.96$
	Noun	$P(\text{verb}   \text{Noun}) = 0.45$	$P(\text{Noun}   \text{Noun}) = 0.05$
	Adj	$P(\text{verb}   \text{Adj}) = 0.02$	$P(\text{Noun}   \text{Adj}) = 0.95$

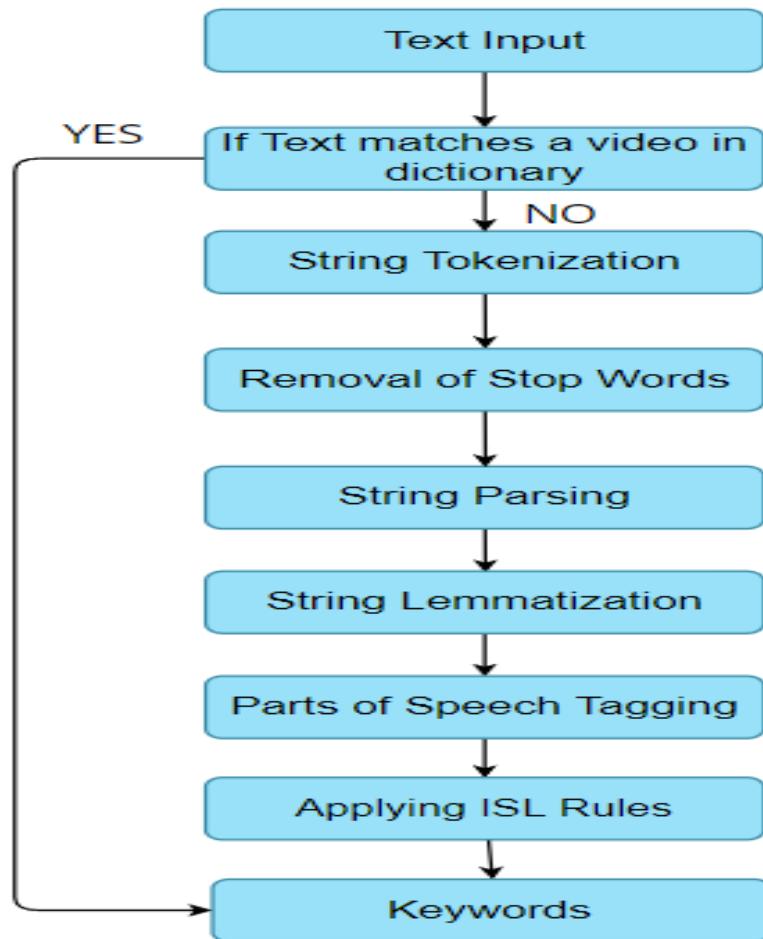
HMM tagging, also known as sequence labeling, is a method of mapping the tag sequence to the input sequence. Assume there is a tagging process with the following inputs:  $X_1, X_2, X_3, \dots, X_m$ . The output will be a tag sequence or state sequence with the following inputs:  $Y_1, Y_2, Y_3, \dots, Y_m$ . The sentence will be the input for part-of-speech tagging, and the tag for each word in the sentence will be the output. As an example, if we have a five-word sentence, the output will be five tags, each representing a different portion of the speech. If the input is a sentence in the source language, the label will be a sentence in the target language in machine translation.

There are three methods for dealing with the tagging problem: the first is the rule-based strategy, which relies on the use of predefined rules. However, rule-based systems can have several issues, including grammatical leaks, the inability to list all the rules, and the last issue, which is the variation in the rules over time, place, and a variety of other factors. The statistical-based strategy can also be used to tackle the tagging problem. Furthermore, a statistically based model is reliant on statistics as well as the availability of a trainable and already labeled corpus. Finally, the hybrid model, which includes both rule-based and statistical-based models, is the most common and practical solution to dealing with the tagging problem.

## 7. Methodology

Figure 6 shows the natural language model's architecture. If the text input which is given as input to the model matches a video in the database, then the input is given as the

keyword; otherwise, the input text is processed through the various NLP technologies and then the keywords are generated. Each part of the technology is discussed in more detail in the heading “Methodology”.



**Figure 6.** Natural Language Model Architecture.

NLTK is the heart of the audio to Indian Sign Language conversion system, as it is the most powerful open-source NLP library which is used to assist with human language data. Text processing is performed using NLTK, which involves various steps, such as tokenization, the removal of stop words, lemmatization, parse tree generation, part-of-speech (POS) tagging, etc.

- **Tokenization**

Tokenization is the process of splitting text into a list of words also known as tokens. NLTK has a module named `tokenize()` which is further divided into two types, i.e., `word_tokenize` to split a sentence into a sequence of words and `sentence_tokenize` to split a paragraph into a list of sentences.

- **The removal of stop words**

Stop words are a list of very common but less informative words that can be ignored. For example—her, me, has, itself, he, so, too, they, them, etc. Since they are not so important in the sentence, they can be removed from the sentence. This will improve the overall performance of the system.

- **Parsing**

Parsing is the syntax analysis phase in which it checks whether the string obtained after tokenization belongs to proper grammar or not. Parsing helps to adjust the text

based on the target language's grammar structure. One of the most used parsers is the Stanford Parser.

- **Lemmatization**

The process of transforming the inflected forms of a word to its root-based dictionary form is known as lemmatization. This root-based dictionary form of a word is referred to as a lemma. The importance of this step lies in ISL, as it requires a root word.

To check the results of lemmatization, we analyze its results through sample sentences. For example—"He was playing and eating at the same time".

The results in Table 5 show that lemmatization alone is not sufficient enough to give accurate root words, as it does not take into consideration the context of the sentence. To overcome this problem, part-of-speech tagging comes into the picture.

**Table 5.** Word-to-root-word conversion using lemmatization.

Word	Lemma
He	He
was	was
playing	playing
and	and
eating	eating
at	at
same	same
time	time

- **Part-of-speech tagging**

POS tagging refers to the process of labeling words with different constructs of English grammar, such as adverbs, adjectives, nouns, verbs, prepositions, etc. POS is a collection of a list of tuples where the first part of the tuple is the word itself and the second part is a tag that identifies whether the word is an adjective, verb, noun, etc.

To check whether part-of-speech tagging can improve the results obtained after lemmatization, we analyze the same sample sentence used above in lemmatization.

The results in Table 6 show that after integrating part-of-speech tagging with lemmatization it gives a correct base form of a word, which in turn improves the accuracy of word to base word conversion.

**Table 6.** Word to root word conversion using POS tagging and lemmatization.

Word	Lemma (after Integrating POS Tags)
He	He
was	be
playing	play
and	and
eating	eat
at	at
same	same
time	time

Thus, the combination of part-of-speech tagging and lemmatization is used in our proposed system to enhance the accuracy of our system.

## 8. Performance Evaluation

The performance of our system is measured based on the net promoter score system. The net promoter score measures the willingness of customers to recommend products or services to their friends and family. A survey was taken with 30 disabled people that asked them to give a rating to the system in terms of how likely they were going to recommend our system to their friend or family who is hearing-impaired.

- Promoters: responses from 9 to 10.
- Passives: responses from 7 to 8.
- Detractors: responses from 0 to 6.
- Total number of people who participated in the survey: 30.
- Total number of promoters: 26.
- Percentage of promoters: 86.6%.
- Total number of passives: 3.
- Percentage of passives: 10%.
- Total number of detractors: 1.
- Percentage of detractors: 3.33%.
- Net promoter score = total percentage of promoters – total percentage of detractors.
- Net promoter score =  $86.6 - 3.33 = 83.27 \sim 83$ .
- A net promoter score above 50 is considered excellent by the creators of the NPS score.

## 9. Results and Discussion

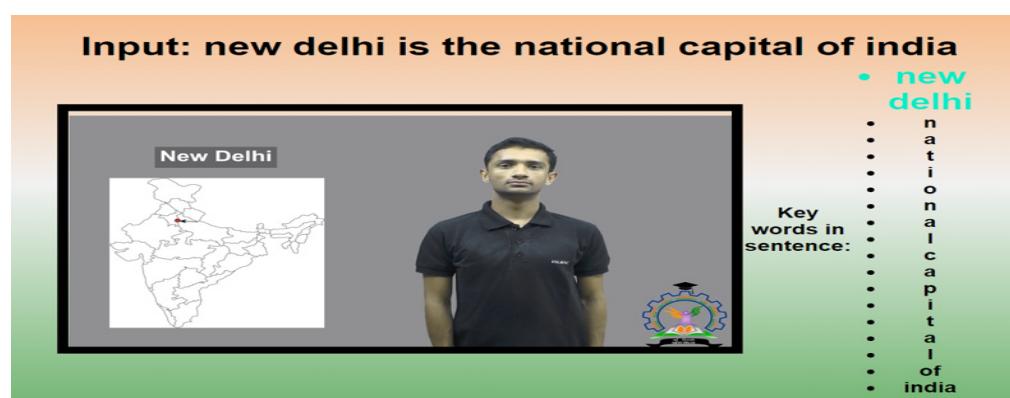
### Case 1:

Input: New Delhi is the national capital of India

ISL sentence: New Delhi national capital of India

Videos shown: {New Delhi, n,a,t,i,o,n,a,l, c,a,p,i,t,a,l, of, India}

In Case 1 (as shown in Figures 7 and 8), the user enters the sentence “New Delhi is the national capital of India”. The sentence/keywords after applying NLP and ISL grammar rules is “New Delhi national capital of India”. In this, “is” is removed, as it is a stop word. The following videos are shown to the user.



**Figure 7.** Case 1. Available Video for ‘New Delhi’.

1. “New Delhi” is shown in one video, as the video for the same is present in the database. This is an example of how the system identifies multiple words/phrases in the sentence for which videos are present in the database.
2. The keywords “national” and “capital” are broken down into letters and a sign language video for each letter is shown to the user as the video, for neither “national” nor “capital” is present in the database.
3. The videos for keywords “of” and “India” are shown to the user, as the sign language videos for both are present in the database and there is no need to further break them into letters.



**Figure 8.** Case 1. Available ISL video for ‘national’ or ‘capital’ or ‘India’.

### Case 2

Input: Change in Temperature

Videos shown: {Change in Temperature}

In Case 2 (as shown in Figure 9), the user enters “Change in Temperature” as the input. The video for the entire input is present in the database and therefore only the video which depicts “Change in Temperature” in sign language is shown to the user.



**Figure 9.** Case 2. Available video for ‘change in temperature’.

### Case 3

Input: Teacher

Videos shown: {t,e,a,c,h,e,r}

In Case 3 (as shown in Figures 10 and 11), the user enters “Teacher” as the input. There is no video present in the database for “Teacher”; therefore, the system breaks the input into letters and shows the videos for each individual letter.



**Figure 10.** Case 3. Available video for letter ‘t’.



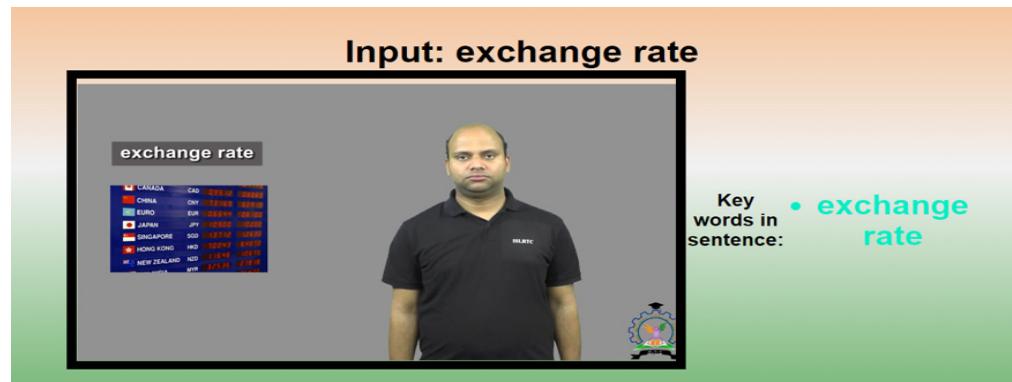
**Figure 11.** Case 3. Available video for letter 'a'.

#### Case 4

Input: exchange rate

Videos shown: {exchange rate}

In Case 4 (as shown in Figure 12), the user enters “exchange rate” as the input. The video for the entire input is present in the database and therefore only the video which depicts “exchange rate” in sign language is shown to the user.



**Figure 12.** Case 4. Available video for ‘exchange rate’.

#### Case5

Input: Kangaroo is an animal

Videos shown: {Kangaroo, a,n,i,m,a,l}

In Case 5 (as shown in Figures 13 and 14), the user enters “Kangaroo is an animal”. The sentence/keywords after applying NLP and ISL grammar rules is “Kangaroo animal”. In this, “is” is removed, as it is a stop word. The following videos are shown to the user.



**Figure 13.** Case 5. Available video for ‘kangaroo’.



**Figure 14.** Case 5. Available video for letter ‘a’.

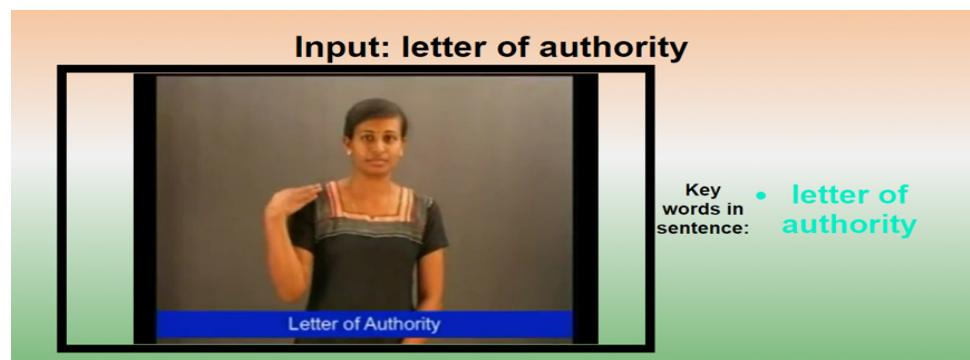
- “Kangaroo” is shown in one video, as the sign language video for the entire word is present in the database.
- “animal” is broken into letters and videos for individual letters which are shown, as the sign language video for “animal” is not present in the database.

#### Case 6

Input: Letter of authority

Videos shown: {Letter of authority}

In Case 6 (as shown in Figure 15), the user enters “Letter of authority” as the input. The video for the entire input is present in the database and therefore only the video which depicts “Letter of authority” in sign language is shown to the user.



**Figure 15.** Case 6. Available video for ‘letter of authority’.

#### Case 7

Input: 2

Videos shown: {2}

In Case 7 (as shown in Figure 16), the user enters the number “2” as the input. Since the numbers between 0 and 9 are present in the database, only one video which depicts “2” in sign language is shown to the user.



**Figure 16.** Case 7. Available video for number ‘2’.

**Case 8**

Input: 30

Videos shown: {3, 0}

In Case 8 (as shown in Figures 17 and 18), the user enters the number “30” as the input. There is no video present in the database for the number “30”; thus, the system breaks the input into two components, i.e., “3” and “0”, and shows the separate videos for both of them.



**Figure 17.** Case 8. Available video for number ‘3’.



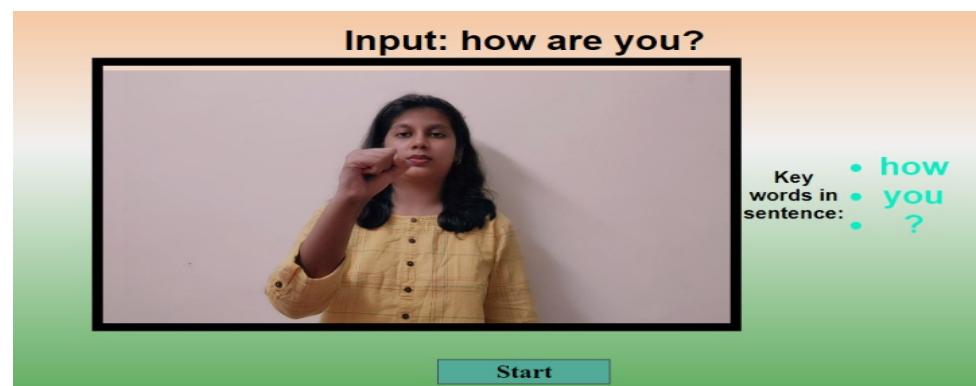
**Figure 18.** Case 8. Available video for number ‘0’.

**Case 9**

Input: How are you?

Videos shown: {How, you?}

In Case 9 (as shown in Figures 19 and 20), the user enters “How are you”. The sentence/keywords after applying NLP and ISL grammar rules is “How you”. In this, “are” is removed, as it is a stop word. Since there is no video present in the database for “How are you”, the sentence is broken into words, i.e., “how” and “you”.



**Figure 19.** Case 9. Available video for ‘how’.



**Figure 20.** Case 9. Available video for ‘you’.

#### Case 10

Input: Good Evening

Videos shown: {Good Evening}

In Case 10 (as shown in Figure 21), the user enters “Good Evening” as the input. The video for the entire input is present in the database and therefore only the video which depicts “Good Evening” in sign language is shown to the user.



**Figure 21.** Case 10. Available video for ‘good evening’.

#### 10. Conclusions

Through this paper, we have presented a user-friendly audio/text to Indian Sign Language translation system specially developed for the hearing- and speaking-impaired community of India. The main aim of the system is to bring a feeling of inclusion among the hearing-impaired community in society. The system does not only help the person who is suffering from a disability but would also be beneficial for the hearing people who want to understand the sign language of a hearing-impaired person so that they can communicate with them in their language. The core of the system is based on natural language processing and Indian Sign Language grammar rules. The integration of this system in areas such as hospitals, buses, railway stations, post offices, and even in video conferencing applications, etc., could soon be proved a boon for the hearing-impaired community in India.

In the future, the features of the system could be enhanced by integrating reverse functionality, i.e., an Indian Sign Language to audio/text translation system which could open the path for a two-way communication system. In addition, the database of the system could be increased to enhance the literacy and scope of the system.

**Author Contributions:** Conceptualization: D.T. and P.S. (Pratibha Sharma); Methodology: P.S. (Purushottam Sharma); Formal analysis and investigation: P.S. (Purushottam Sharma); Writing—original draft preparation: D.T. and N.N.; Writing—review and editing: C.V. (Chaman Verma); Funding

acquisition: none; Resources: P.S. (Purushottam Sharma); Supervision: P.S. (Purushottam Sharma). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hanke, T. HamNoSys—Representing Sign Language Data in Language Resources and Language Processing Contexts. *LREC 2004*, 4, 1–6.
2. Prikhodko, A.; Grif, M.; Bakaev, M. Sign Language Recognition Based on Notations and Neural Networks. In Proceedings of the Communications in Computer and Information Science, Valletta, Malta, 25–27 February 2020; Volume 1242, pp. 463–478. [[CrossRef](#)]
3. Sonawane, P.; Shah, K.; Patel, P.; Shah, S.; Shah, J. Speech to Indian Sign Language (ISL) Translation System. In Proceedings of the IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS, Greater Noida, India, 4–5 November 2021; pp. 92–96. [[CrossRef](#)]
4. Tewari, Y.; Soni, P.; Singh, S.; Turlapati, M.S.; Bhuvan, A. Real Time Sign Language Recognition Framework for Two Way Communication. In Proceedings of the International Conference on Communication, Information and Computing Technology, ICCICT, Mumbai, India, 25–27 June 2021. [[CrossRef](#)]
5. Kunjumon, J.; Megalingam, R.K. Hand gesture recognition system for translating Indian sign language into text and speech. In Proceedings of the 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 27–29 November 2019; pp. 14–18.
6. Gangadia, D.; Chamaria, V.; Doshi, V.; Gandhi, J. Indian Sign Language Interpretation and Sentence Formation. In Proceedings of the 2020 IEEE Pune Section International Conference, PuneCon 2020, Pune, India, 16–18 December 2020; pp. 71–76. [[CrossRef](#)]
7. Shangeetha, R.K.; Valliammai, V.; Padmavathi, S. Computer vision based approach for Indian sign language character recognition. In Proceedings of the 2012 International Conference on Machine Vision and Image Processing, MVIP, Coimbatore, India, 14–15 December 2012; pp. 181–184. [[CrossRef](#)]
8. Sawant, S.N.; Kumbhar, M.S. Real time sign language recognition using PCA. In Proceedings of the 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT 2014, Ramanathapuram, India, 8–10 May 2014; pp. 1412–1415. [[CrossRef](#)]
9. Papadogiorgaki, M.; Grammalidis, N.; Tzovaras, D.; Strintzis, M.G. Text-to-sign language synthesis tool; Text-to-sign language synthesis tool. In Proceedings of the 2005 13th European Signal Processing Conference, Antalya, Turkey, 4–8 September 2005.
10. Sonare, B.; Padgal, A.; Gaikwad, Y.; Patil, A. Video-based sign language translation system using machine learning. In Proceedings of the 2021 2nd International Conference for Emerging Technology, INCET, Belagavi, India, 21 May 2021. [[CrossRef](#)]
11. Kanvinde, A.; Revadekar, A.; Tamse, M.; Kalbande, D.R.; Bakereywala, N. Bidirectional Sign Language Translation. In Proceedings of the International Conference on Communication, Information and Computing Technology, ICCICT, Mumbai, India, 25–27 June 2021. [[CrossRef](#)]
12. Qi, J.; Wang, D.; Jiang, Y.; Liu, R. Auditory features based on gammatone filters for robust speech recognition. In Proceedings of the 2013 IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013.
13. Nair, M.S.; Nimutha, A.P.; Idicula, S.M. Conversion of Malayalam text to Indian sign language using synthetic animation. In Proceedings of the 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), Kottayam, India, 1–3 September 2016.
14. Qi, J.; Wang, D.; Xu, J.; Tejedor Noguerales, J. Bottleneck features based on gammatone frequency cepstral coefficients. In Proceedings of the International Speech Communication Association, Lyon, France, 5–29 August 2013.
15. Grif, M.; Manueva, Y. Semantic analyses of text to translate to Russian sign language. In Proceedings of the 2016 11th Inter-national Forum on Strategic Technology, IFOST, Novosibirsk, Russia, 1–3 June 2016; pp. 286–289. [[CrossRef](#)]
16. Dhanjal, A.S.; Singh, W. Comparative Analysis of Sign Language Notation Systems for Indian Sign Language. In Proceedings of the 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), Sikkim, India, 25–28 February 2019.
17. Varghese, M.; Nambiar, S.K. English to SiGML conversion for sign language generation. In Proceedings of the 2018 Interna-tional Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 21–22 December 2018.
18. Raghavan, R.J.; Prasad, K.A.; Muraleedharan, R. Animation System for Indian Sign Language Communication using LOTS Notation. In Proceedings of the 2013 International Conference on Emerging Trends in Communication, Control, Signal Processing and Computing Applications (C2SPCA), Bangalore, India, 10–11 October 2013.
19. Dhivyasri, S.; Krishnaa Hari, K.B.; Akash, M.; Sona, M.; Divyapriya, S.; Krishnaveni, V. An efficient approach for interpretation of Indian sign language using machine learning. In Proceedings of the 2021 3rd International Conference on Signal Processing and Communication, ICSPC, Coimbatore, India, 13–14 May 2021; pp. 130–133. [[CrossRef](#)]

20. Allen, J.M.; Foulds, R.A. An approach to animating sign language: A spoken english to sign english translator system. In Proceedings of the IEEE Annual Northeast Bioengineering Conference, NEBEC, Troy, NY, USA, 7–19 April 2015; Volume 30, pp. 43–44. [[CrossRef](#)]
21. Patel, B.D.; Patel, H.B.; Khanvilkar, M.A.; Patel, N.R.; Akilan, T. ES2ISL: An Advancement in Speech to Sign Language Translation using 3D Avatar Animator. In Proceedings of the Canadian Conference on Electrical and Computer Engineering, London, ON, Canada, 30 August–2 September 2020. [[CrossRef](#)]
22. Agarwal, A. Generating HamNoSys signs from the user’s input. In Proceedings of the 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India, 4–5 September 2015.
23. Priya, L.; Sathya, A.; Raja, S.K.S. Indian and English Language to Sign Language Translator—an Automated Portable Two Way Communicator for Bridging Normal and Deprived Ones. In Proceedings of the ICPECTS 2020—IEEE 2nd International Conference on Power, Energy, Control and Transmission Systems, Proceedings, Chennai, India, 10 December 2020. [[CrossRef](#)]
24. Saija, K.; Sangeetha, S.; Shah, V. WordNet Based Sign Language MachineTranslation: From English Voice to ISL Gloss. In Proceedings of the 2019 IEEE 16th India Council International Conference (INDICON), Rajkot, India, 13–15 December 2019.
25. Ahire, P.G.; Tilekar, K.B.; Jawake, T.A.; Warale, P.B. Two-way communicator between deaf and dumb people and normal people. In Proceedings of the 1st International Conference on Computing, Communication, Control and Automation, IC-CUBEA 2015, Pune, India, 26–27 February 2015; pp. 641–644. [[CrossRef](#)]
26. Jamil, T. Design and Implementation of an Intelligent System to translate Arabic Text into Arabic Sign Language. In Proceedings of the Canadian Conference on Electrical and Computer Engineering, London, ON, Canada, 30 August–2 September 2020. [[CrossRef](#)]
27. Suleiman, D.; Awajan, A.; Al Etaiwi, W. The Use of Hidden Markov Model in Natural ARABIC Language Processing: A survey. *Procedia Comput. Sci.* **2017**, *113*, 240–247. [[CrossRef](#)]