International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

ISL Smart Translator: Speech and Text to Indian Sign Language Converter

Miss. Urvi Pawar, Miss. Shreya Rakibe, Miss. Vaishnavi Sandhan,
Miss. Samruddhi Vispute, Assistant Professor Dr. Kirti Patil
Dept. of Information Technology MET's BKC
Adgaon, Nashik, India

Abstract- Communicating with Deaf or hard-of- hearing people can be difficult at times and the chal- lenges are great in multilingual countries such as India. This study describes current work on this but offers a proposal for an animated ISL (Indian Sign Language) translation system from Marathi text and/or speech – based on the fact that there are many more English- MSL (Marathi Sign Language) resources available and, therefore, a 'significant access gap' when considering Deaf users within our target country. The majority of currently available translation systems have been based upon machine learning; however, due to insuffi- cient parallel corpora/annotated sign data resources for Marathi-MSL, this method will not work. The proposed system adopts as an alternative a 'rule based' method- ology which will map the Marathi language & grammar structures onto ISL using linguistic 'rules' and dictio- nary and will develop this through web application using React.jsTailwindCSSFlask for the front end of the web application, while allowing use of 'browser based' storage thus ensuring a very lightweight deployment. Given that ISL requires 'gestures', 'facial expressions' and 'spatial syntax', it is not possible to translate word for word; rather, the system will also consider some important Marathi grammatical elements, e.g. inflection; location/post position, verb forms, sentence structure and will therefore generate an ISL output that more accurately represents the original Marathi and promotes communication and accessibility for Deaf individuals.

Keywords— Marathi to ISL translation, Indian Sign Language, Rule-based machine translation, Low resource language processing, Web-based sign language generation, Accessibility technology, regional language NLP, Dictionary-based translation, Simultaneous mor- phology, Inclusive communication systems.

I. INTRODUCTION

Communication with others and learning through lan- guage is critically important for humans, yet it creates bar- riers for individuals with hearing impairments. Most class- rooms rely on spoken and written communication, making it difficult for sign language users to fully understand or participate. Although sign language is a complete and effective form of communication, it is not widely known among educators and students. This lack of awareness lim- its inclusion and reduces access to mainstream education. As a result, students with hearing impairments often face challenges in comprehension and classroom engagement. Bridging the gap between spoken and sign language is essential to ensure equal learning opportunities. Previous research has developed translation systems similar to ASL to English, but these mainly support languages with large datasets. Machine learning methods require extensive par- allel data, which is not available for regional languages like Marathi. This creates a significant gap in assistive technology for low-resource languages. Additionally, many systems function as "black boxes," making their processes unclear and difficult to trust. This lack of transparency is problematic in education, where accuracy is critical.

II. LITERATURE SURVEY: INFLECTION RULES FOR MARATHI TO ENGLISH IN RULE BASED MACHINE TRANSLATION

1. Domain overview and historical evolution

This proposed project looks at connecting the ap- proaches of spellout and rule base methods for the con- struction of Marathi words, which will allow for English to be produced from the base letters of all complex Marathi words. This project will also use the same technique as was used in the above study by having subject matter experts define the rules of grammar rather than relying on a statistical model to do this. The continued applicability of the above noted method is extremely valuable when compared to the rigidities of rule based systems (Marathi will have a different order for the placement of words in comparison to English). In Marathi, all verbs appear after the subject, but in English the second placement of the verb is relative to other verbs; there can be differences in where some words occur at the end of a sentence or where a noun occurs at the start of a sentence.

2. Detailed analysis of the base paper
An acceptable basis for translating into English may be provided through the use of the inflectional patterns of four major classes of Marathi words, namely nouns, pro- nouns, verbs, and adjectives; as well as with case/position (changes) from the four major types of verbs.

3. Methodological Approach and Mathematical Formulation
Separating tokens (step one) will allow you to use shallow parsing (as in breaking apart the way something appears) to do the morphological analysis to suggest some possible inherent forms for the tokens or surfaces that were generated as a result of tokenizing the original content without using statistical methods. As a result, the inflections of the newly generated tokens will be based solely on the use of a small number of standard tables and an associated rule set (grammar) – which means that there are no ”predictions” of how the SOV word order will eventually end up as SVO, but instead, the SOV and SVO must be built sequentially by building up the simpler elements to the more complex until an SVO configuration is reached. Finally, you have a set of new English words instead of an equation, random answer/object). Because of this, the explicit manner in which the patterns create the new words cannot be inferred directly from the underlying data.

4. Experimental Validation and Performance Results
From TDIL Corpus data, we found 25,000 sets of Marathi and English Translated Tourism Phrase Sets. Based on how well they did (overall) 7,000 were chosen as valid tests of Performance with Average Accuracies of 88%–90%. The average precision ratings were virtually the same as the overall performance acc. ones therefore they are considered a measure of the system’s overall quality.

5. Practical Contributions and Strengths
There is a clear listing of inflectional patterns and clear mappings from Vibhakti to prepositions, each output of which is created via a set of open rules. The reproduction of the tool can be made easy by virtue of the tool being part of IIIT Hyderabad’s common shallow parser, which also allows the tool to be adapted more widely in Frameworks for Indian languages.

6. Limitations
Aside from typical tourist site translations, you don’t have any assurance about getting machine Translated con- tent — which is your challenge with self-providing a guar- antee. But with the current way of performance evaluation (BLEU, METEOR) your tools cannot effectively assess your output because there was no available performance measures. This also will be compounded because with the incomplete definitions of translation errors and having no clear way to identify the translation errors, you cannot identify which parts of the translated text contributed to the error. This may lead to your solutions not being able to recreate the basic characteristics or signs of a sign language (e.g., their meaning, cultural expression and their relationship to concepts).

III. LITERATURE SURVEY: MARATHI TO INDIAN SIGN LANGUAGE MACHINE TRANSLATION

1. Domain Description and Historical Development
The focus for research into speech-to-sign machine translation has traditionally been on natural language processing, tools for accessing the translation, and finally the structure of the language to be translated. Research has progressed towards more complex methods for trans- lating spoken languages into sign languages since 1990. The progress has moved from translating single words to translating sentences, translating sentence structures, using probabilistic modelling and recently to using deep learning. However, one of the main pitfalls of contem- porary approaches is due to the multi-layered, or com- pounded, structure of Indian Sign Language; where one concept/meaning is represented at the same time by hand- shape, facial expression and body position, rather separate from one another.

2. Objectives and Key Aspects of the Paper
The objective of producing a quality Marathi-to-Indian Sign Language translation system will be reached with the help of gloss outputs produced in conjunction with live performance by an interpreter. Unlike many previous systems (such as BERT), which rely upon large numbers of data sources, this system will produce output based upon rules established for linguistics. There are two primary questions that will be answered through the research project: How does ISL allow multiple grammatical changes to occur simultaneously within a sign. The second question is how will the symbolic notation created through Ham- NoSys and the SiGML encoding schemas be converted into movement. Assessment will occur on three dimensions— total clarity of generated output; total smoothness of generated output; and accuracy of generated output.

3. Methodology and Mathematical Formulation

The process begins with tokenization, followed by lemmatization. The next step involves assigning parts of speech before proceeding to shallow parsing. Then an
analysis of morphology and syntax occurs before creating glosses. Finally, the final element of this entire process is animation through SiGML. One main algorithm governs how signs change related to tense, negation, or questions. The other adjusts visual parameters such as hand position, movement, and mouth shape while in real time as you sign.

4. Experimental Validation and Performance Results

The basis of this project was based on an initial core of over nine hundred pairs of words from Marathi to In- dian Sign Language. The study employed two professional sign language interpreters and thirteen educators trained to evaluate performance based upon over seven hundred testing case methods. The average score for each testing case was 4.3 out of five for maintaining the meaning of the source signs in both ascending and descending order. The average score for fluency or appearance of the output was 4.1. The average accurate results for each syntactic type fall between 3.8 to 4.8.

5. Strengths and Practical Contributions

The creation of a new structured system for linking the creation of a gloss in Marathi-ISL with an animated visual representation of that gloss via motion is a significant advance in the use of technology for sign communication. This new approach overcomes the limitations of previ- ous attempts by processing grammatical structures with overlaps in real time using a detailed part-of-speech label based on the differences between spoken Marathi and signed ISL. The produced animation will utilize standard HamNoSys and SiGML code so that the resultant signing sequence appears in 2-dimensional space with a clear two- dimensional video representation for each signing sequence or gloss. This new methodology will have significant value for everyday signing interaction and in educational con- texts as well.

6. Limitations

The system does not warn users when it discards an unknown word, even though it is able to process simple sentence structures of a certain type. However, using the same verb repeatedly to indicate aspect is not normal, and only the interrogative form uses non-hand signals (e.g., through facial expressions). Complex language systems will lose their established rigidity as they develop complex- ity; therefore the system will find it increasingly difficult to work with larger structures when they appear.

IV. SUPPORTING LITERATURE AND RESEARCH INTEGRATION

1. Statutory Approaches to develop machine translation systems for the Marathi language

Since the majority of news articles concerning machine translation do not mention anything about Marathi due to the lack of data available from both large technology companies and/or colleges/universities, there are very few
sources containing bilingual documents for referring to. Thus, rule-based systems remain the only relative form of viable machine translation into the Marathi language due to insufficient bilingual document reference sources providing sufficient corpora for machine training. Simple word substitution in English into Marathi does not pro- duce an acceptable translated product because English has a subject-verb-object (SVO) sequence while Marathi uses a subject-object-verb (SOV) sequence. In addition, many temporal indicators present in English will have no correlation with those same temporal indicators and/or tenses used in the Marathi language – thus, creating an abundance of mismatch errors associated with temporal indicators within the machine translation output. Further- more, when translating into Marathi, placing adjectives may also produce an abundance of mismatch errors since no corresponding type of placement exists within the structure of both languages.

2. Sign Language Machine Translation: Architectural Insights

ISL develops its meaning all at once through hand shapes and other signs along with facial expressions and body postures. There is no tense based on the use of verb forms, with tense existing as a singular entity (e.g., 'run' versus 'ran') or separately from the original form. Pronouns are not determined in relation to gender, and there is no grammar-based use of pronouns as is seen in many spoken languages. Many concepts may also develop from signing a repeated sign instead of from a plural form. A question may be recognized by lifting your eyebrows, not by altering word order. The use of groups of signs studied by Bhagwat's group produced more easily understood structures than the use of analysing individual items only.

3. Transfer Learning and Cross-Lingual Approaches

Despite being populated by only thousands of examples an effort involving under-resourced Arabic has reached 80% accuracy indicating that such as a rule can hold up well over time. It was found through research that having human-created high quality inputs performed better than having large amounts of machine-generated data which typically contain many

errors. Models built to connect two separate languages performed better than those designed to work on multiple languages making it logical to focus future efforts for Marathi-ISL based on the present out- come ratings obtained from the two language development efforts.

4. Dealing with Language Differences

Some significant challenges that arise during the trans- lation process are: To perform a proper translation from the accusative case 'ला' we must use semantic mapping e.g., 'मुलाला आंबा हवा आहे' (A boy wants a mango). Another significant challenge to making a proper translation of the instrumental case 'ने' is using a contextual transformation to translate an English example such as 'रामने पत्र िलहले' (Ram wrote a letter) since there is no subject case in English. If there are cultural concepts that exist in one language but not another, such as कमर् (action) and धमर् (duty), they will not have a direct translation into English. Furthermore there many time expressions in Marathi e.g., सकाळ (morning), दुपार (afternoon), संध्या (evening) and रात (night) which do not have an equivalent time expression in English so the time mapping will be different from English to Indian Sign Language (ISL). Also, using reduplication to convey more than one aspectual meaning e.g. using the word 'हळू हळू' (slowly) and 'पाहता पाहता' (as you are looking), several ISL signs will be needed to represent these words.

5. Multimodal Output Generation

IESE chooses pre-recording video clips instead of Ham- NoSys due to web browser speeds. Facial expressions, i.e. raised eyebrows and tilted heads, become included in the visual clip files according to required Indian Sign Language (ISL) grammar. Useability of pre-recording video clips pro- vides simple implementation in comparison to animated files. Ultimately, there is a balance between accuracy of ISL and practical delivery online.

6. Rule-Based Translation Validation

Over a training dataset size of 10,000 pairs, rule-based systems perform better than neural systems; for instance, the Marathi-to-ISL (Indian Sign Language) dataset con- sists of less than 200 sentence pairs. Most common sen- tence patterns can be identified by initial sets of rules; for rarer patterns, additional rules will be needed in order to create a properly functioning system. Since metric systems such as BLEU and METEOR are ineffective when they are used on sign language translations, a third person must evaluate accuracy. In Bhagwat's experiment, he tested over 700 sentence pairs and received an overall rating (totaling both adequacy and fluency) of 4.30 and 4.10 respectively.

V. PROPOSED SYSTEM INTEGRATION ARCHITECTURE

1. Proposed System Architecture

- Modular System Elements: The System contains 4 key components: Natural Language Processing (NLP), Rule Based Machine Translation (RBMT), Gesture Mapping, and Gesture Rendering via browser.
- Input Processing: Input to the system may be text or speech and speech will be converted into text before processing through NLP techniques including, but not limited to, tokenizer, normalization, morpho- logical analysis, part of speech tagging, and named entity recognition.
- Rule Based Transfer Translation and Gesture Map- ping: Convert the structure of the Marathi sentence (SOV) to that of an English-like structure (SVO), in a rule- based transfer approach for handling the tense, negation, question formation, and ordering of the sentences. Export from RBMT is the simplifi- cation of the English Gloss, to generate appropriate gestures -mapping between the latter and the actual gesture for ISL.
- Back End Implementation: The back-end of the system is implemented with a Flask API to manage tokens, grammar rules, gesture positions, caching, and overall performance of the system.

2. Data Flow and System Process

- User Input and Speech-to-Text Conversion: User are able to put in Marathi text or convert Marathi speech to text for processing
- Text Normalization: The enough text is cleaned and normalized with the intent of changing the unstruc- tured form to a more structured form that is more easily processed
- Linguistic Analysis By NLP: The text undergoes to-kenization, morphological analysis, tagging of part- of-speech (POS) and tagging of named (entities) in order to identify the semantics and grammatical structures of the words.
- Linguistic Annotation of Output: The annotated tokens consisting of part-of-speech (POS) tags (mor- phemes), (possible) morphological features and (the actual) syntactic relationships will be used to assist in translating into ISL.
- Rule Based Translation Process: The annotated data will be processed through the Rule-Based Transfer module (3 stages):
- Analysis
- Transfer of data into a new form
- Generating new ISL text

3. Use Cases and Operational Scenarios

- Teaching Support for Inclusiveness in the Classroom: The Marathi-to-ISL translation system will convert Marathi commands issued to Deaf students into ISL hand gestures in real-time allowing teachers to communicate with both hearing and Deaf students at the same time.
- Classroom Instruction Translated in Real Time: For example, common classroom commands such "पुस्तक उघडा," "गृहपाठ पूणर् करा," and "प्रश्न सोडवा" will now be visually illustrated in ISL hand gestures for deaf students, allowing them to understand the command immediately.
- Daily Communication Support for Schools: For ex- ample, if a Deaf student inputs a Marathi sentence such as "मला पाणी पाि◌हजे.", "आज सुट्टी आहे का?", or "तुम्ही मला मदत कराल का?" into the system, he or she will receive ISL hand gestures in response for communi- cating with teachers and classmates.
- Communication Without Interpreter: The system allows for seamless communication between Deaf students and their teachers and classmates without the need for a human interpreter in day-to-day school interactions.
- ISL Self-Study Tool: The system can also provide a self-study tool with an illustrated (virtual) dictio- nary of words in ISL, grammar, sentences, and facial expressions. The self-learning tool will allow users to learn ISL words by way of comparing the image of the hand movements in ISL with the text descriptor of the English words and also by way of comparing other users' videos to see how the correct gesture was performed.

VI. TECHNICAL COMPARISON AND INTEGRATION FEASIBILITY

1. Architectural Compatibility Assessment

The proposed solution has a strong correlation with both referenced papers. The NLP preprocessing loop will closely mimic Anselmo's RBMT workflow (Tokenization → Morphological Analysis → POS Tagging → Shallow Parsing → Syntactic Reordering). Hence, integration of inflection and transformation rules will require minimal adjustments to be accomplished. The module structure is similar to that of the four stage Marathi-ISL MT paper (Rule-Based Transfer Translation → Gesture Mapping → Front End Visualization). Both of the referenced pa- pers utilize a similar method of treatment for Marathi case markers (-ला, -ने, -मध्ये), morphological structures, and phrase-based shallow parsing. Therefore, rule sets, inflec- tion tables and gesture sequencing strategies from the referenced papers can all be incorporated directly into the proposed system without changing the structure.

2. Computational Requirements and Deployment Analysis

This setup runs efficiently on simple devices from a technical standpoint. There is no need for power-hungry processor designs because the processes use simple rules - separating words into parts and identifying their type. In- stead of creating and displaying signs in real-time through complex coding processes, only animated images of video clips are used to produce the motion/animation of the sign being displayed. The actual size of animated video clips is very small. Because they are directly stored within a web browser, the amount of data sent with the video clips when they are viewed significantly decreases the amount of data sent over a thin network. As a result, erratic connectivity in remote locations is less of an obstacle. The areas where efficiency matters are clear. One web link will open the Flask and React set up, and it works on every type of device from tablets to desktop computers, and it can be hosted by any budget-friendly or major cloud provider. Every time guidelines or word lists change, each item can be previously established functionality. In addition, function- ality will remain consistent, even as more functionality is added, to meet the ever-increasing demand.

3. Performance Expectations and Operational Characteristics

Straightforward phrases used in the classroom are the clearest indicators of accuracy: that's what we want to accomplish. It only takes 1 second or less from the time a student hears a Marathi word to the time it turns into an ISL gesture because all of the matching procedures are fixed. Students and teachers can use video snippets at any speed because they are prepared in advance; slow internet connections or old computers will not impede any student or teacher from performing their duties. Occasionally, a teacher or student may use a word that is difficult to express, but the decision making processes are clear so the results can be anticipated. Deaf teachers and students will use outputs more readily because they understand how each of the steps works. Expanding vocabulary or gram- mar will not affect timing because the patterns determine matching; not the amount of training data.

VII. IMPLEMENTATION AND DEPLOYMENT CONSIDERATIONS

1. Database/System Storage Design

The design of our storage system is modular and lightweight, separating linguistic resources and multime- dia gesture assets. The linguistic resources such as the morphology tables of the Marathi language, POS tag map- pings, and inflection rules,

etc., are stored in a relational or document-oriented database as they are primarily textual and require only small amounts of data storage. ISL ges- ture video and image assets are stored outside the backend database in a media repository or content delivery storage to separate the scalability of the gesture library from the performance of linguistic processing. Cacheable gesture clips, when used frequently, can be cached in the client browser, thus allowing gesture clips to operate offline or in low-connectivity environments while also reducing calls to the backend database. Together, this simple, maintainable storage architecture enables this system to be deployed in an educational setting.

2. Hardware Deployment Options

Depending on your environment, the system provides multiple possibilities for hardware deployment. In cen- tralized environments (like schools) the backend can be hosted on a small local server or institutional PC, running via the Flask APIs with minimal load on the server. The system does not require GPUs; therefore, it can easily run on standard CPUs due to its rule-based architecture. For larger or distributed environments, the backend can be hosted in a cloud system; therefore, multiple Local Service Providers (LSPs) can provide services from one location.

VIII. CONCLUSIONS AND FUTURE DIRECTIONS

1. Key Findings from Literature Analysis

Issues with Structural/Grammatical Differences and Due to the Structural Differences between the Two Languages: The two languages have vastly different syntax. While Marathi generally follows a subject- object-verb construction and utilizes case markers or suffixes to indicate the relationships between words in sentences, ISL employs a great deal of flexibility in constructing sentences, relying upon the use of spa- tial grammar, as well as a combination of handshape, movement, and facial expression. Additionally, the fact that the grammar of ISL is constituted simulta- neously by all multiple features of the expression of a single sign makes it impossible to translate word for word and using only word-for-word would result in an unnatural manner of signing.

Issues with Dyslexia Standards (Standards) and Di- alectal Variation: Signed dictionaries that are de- veloped using a standard, such as HamNoSys and SiGML, allow for the proper storage and retrieval of signs. However, the regional differences in Marathi dialects, such as those found in Pune versus Nagpur, Konkan, etc., will also need to be taken into account when performing translations because the variations in dialect will have an impact on the accuracy of the translations and the usability of the translated system.

2. Contributions of This Study

Research Foundation and Gap Identification: This study integrates research from machine translation, sign linguistics, Marathi NLP and accessibility to identify key gaps such as lack of datasets, morphol- ogy processing guidelines, simultaneous grammars, dialect support, and evaluation methods.

System Design and Linguistic Analysis: A realistic dictionary-based system was designed that does not require large datasets, and compares Marathi with other languages in terms of grammar, sentence struc- ture, morphology as well as translation difficulties like word ordering, case-marking, and verb inflection.

Technology Implementation and Development Ap- proach: The development of the system will include the use of React, Flask and browser storage; the study aims to show how linguistic theory can be ap- plied to technological implementation by presenting a phased development procedure ranging from basic vocabulary through refinement via user input.

3. Directions for Future Research

Development of linguistic resources: The future work should involve developing additional gesture dictio- naries; developing a parallel sentence dataset fea- turing Marathi/ISL along with video annotations;
and establishing Marathi morphological databases containing verb conjugation information for different word forms.

Speech Processing and Dialect Support: The system must be able to identify Marathi speech regardless of accent or noise level. The system must maintain context between sentences. The system must be able to identify at least the four main types of Marathi dialects and how each is related to its dialect map.

Evaluation and User Testing: There should be au- tomatic evaluation metrics created. Deaf ISL users should be used for user testing to verify the quality, naturalness and usability of the translation.

REFERENCES

1. A. Alshahrani and H. Alsubaie, “Enabling two-way communica- tion of deaf using Saudi Sign Language,” International Journal Of Advanced Computer Science and Applications (IJACSA), vol. 13, no. 9, pp. 89–98, 2022.

2. S. Sawant and A. Kulkarni, “Marathi To ISL machine transla- tion,” in Proceedings of the 14th International Conference on Natural Language Processing (ICON-2017), pp. 292–301, 2017.
3. S. Alqahtani, M. Elaraby, and M. Diab, “Semantic textual similarity for Modern Standard and Dialectal Arabic using transfer learning,” in Proceedings of the Sixth Arabic Natural Language Processing Workshop, Association for Computational Linguistics, pp. 179–189, 2021.
4. A. Patil and A. Bharati, “Handling challenges in rule-based machine translation from Marathi to English,” International Journal of Computer Applications, vol. 148, no. 13, pp. 1–7, 2016.
5. A. Joshi and H. Barot, “Inflection rules for Marathi to English in rule-based machine translation,” International Journal of Computer Science and Engineering (IJCSE), vol. 7, no. 4, pp. 1070–1075, 2019.
6. S. Sawant and A. Kulkarni, “Marathi–Indian Sign Language bilingual dictionary (prototype resource),” created as part of ICON-2017 system, 2017.
7. Technology Development for Indian Languages (TDIL), “Marathi–English parallel corpus (Tourism domain),” Government of India, Ministry of Electronics and Information Technology (MeitY), 2014.
8. A. Bharati, D. M. Sharma, and R. Sangal, “Shallow Parser Tools for Indian Languages,” IIIT Hyderabad, 2006.
9. A. Bharati and R. Sangal, “Rule-based machine translation for Indian languages: Design and evaluation challenges,” in Proceedings of the IEEE Conference on Natural Language Pro- cessing Systems, 2024.
10. K. V. S. Rao and P. V. Rao, “Grammar-driven machine trans- lation approach for morphologically rich Indian languages,” in Proceedings of the IEEE International Conference on Compu- tational Linguistics, 2025.