Machine Translation of multiple languages into Indian Sign Language Glossary using Dependency Parsing

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science in ***Computational Linguistics*** *by Research*

by

Abhigyan Ghosh
20171089
abhigyan.ghosh@research.iiit.ac.in

International Institute of Information Technology, Hyderabad
Hyderabad 500032, India
July 2023

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Machine Translation of multiple languages into Indian Sign Language Glossary using Dependency Parsing" by Abhigyan Ghosh, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date	Advisor: Prof. Radhika Mamidi

To *IIIT Hyderabad*

Acknowledgments

I would like to thank my family, all of my friends, and all of my Professors, without whose continuous help, support, and guidance, this work would never have been possible.

Abstract

Sign languages have possibly been around for longer than spoken languages. What is interesting about them is that they evolve similarly to spoken languages and tend to have complex grammar and vocabulary. They are mostly used in day-to-day communications by the hearing-impaired community or by people who interact with them on a daily basis. It is even considered by some to be their mother tongue.

Indian Sign Language is used mainly in the Indian subcontinent, mostly by the hearing-impaired community. Apart from investigations that have a more or less sociological approach, research results about the linguistic structure of sign language in India had been very scarce until the last decade. However, it has gained prominence in recent years. The 11th Five Year Plan (2007-2012) acknowledged that the needs of people with hearing disabilities had been relatively neglected and envisaged the development of a sign language research and training center, to promote and develop sign language and training of teachers and interpreters.

As such, some work is being done in translating spoken languages in various forms into ISL. But most of the work done in this regard always focuses on the accuracy of the signed videos while grammar takes a back step. In 2 we show that these tools are mainly used by hearing-impaired students who are trying to pick up English as a second language. Our work hence focuses on getting as much grammatical accuracy as possible for converting English to Indian Sign Language. Here we tackle the problem of converting written text into ISL video by using Text for Indian Sign Language as an intermediary step. To do this, we use rule-based translation methods and apply techniques such as multi-word expression detection and synonym substitution in the source language. In this thesis, I shall try to explain the advantages of taking this route as well as provide some background into the world of sign languages with a focus on Indian Sign Language (ISL).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

According to Census 2001, there are more than one million people with hearing disabilities and more than ten million people are hard of hearing in India. Indian Sign Language is required to give educational, vocational, social, and personal guidance and counseling to such people. English has emerged as the language which is mostly used in these fields but people with hearing disabilities have a hard time understanding regular spoken languages such as English and Hindi as they are educated in and mostly use ISL as the primary language of communication. We hope that a translation system from English could be helpful for such purposes.

Machine Translation between texts in two different spoken languages is a well-defined Natural Language Processing task with approaches ranging from rule-based methods to recent Deep Neural Network Architectures. However, most of the work done in the domain of Sign Language translation seems to be in the domain of Computer Vision and Gesture Recognition fields with very little emphasis on the grammar of ISL []. As such the quality of translations achieved from these systems is very poor as they perform rudimentary sentence transformations. Our aim was to improve the translations by introducing a layer of linguistic manipulation into the translation systems. For this, we formally introduce the vocabulary of Indian Sign Language which we have named Text for ISL as well as translation of any language to Text for ISL.

1.1.1 Importance of Indian Sign Language

Hearing-impaired people use sign language using handshapes, fingers, facial expressions, gestures, and other body parts. It is a visual-spatial language as the signer often uses the 3D space around his body to describe an event. Sign languages, until the 1960s, were not viewed as bona fide languages but just collections of gestures and mimes. Dr Stokoe's research on American Sign Language proved that it is a full-fledged language with its own grammar, syntax, and other linguistic attributes [30]. There are some efforts to prove the same for other sign languages, including Indian Sign Language. [35]

The sign language used in India is Indian Sign Language (ISL). A study by [32] specifies that the ISL used in different parts of India is almost identical in its structure, with differentiation in signs. It is a social need to encourage the hearing-impaired individuals of the Indian sub-continent with a tool that can translate the English text to ISL. [15] reports that 5% of deaf children get educated formally. According to [40], before 1926, there was no formal education available for ISL and different schools used different sign systems.

1.2 Challenges

The biggest challenge we face with such a niche translation task is the lack of publicly available bilingual datasets which is essential for running data-intensive deep learning algorithms. Apart from this, there has been very little linguistic work done on ISL Grammar. There are only two recognized books on ISL Grammar that are used for teaching in various special ability schools in India. Most of the work has been done in formalizing the vocabulary of ISL and manually converting educational material into ISL.

1.3 Problem Formulation

So the problem we are broadly tackling is the conversion of text in any given language into ISL Videos but the core task in this entire approach is the conversion of text into Text for Indian Sign Language. This Text for Indian Sign Language can then be converted into ISL Videos in many ways such as HamNoSys [12] or simply stitching together ISL Videos of the respective signs directly.

1.4 Proposed Approach

We propose a rule-based method for the translation of text in any given language to Text for Indian Sign Language. Traditionally rule-based methods apply rule-based transformations on a constituency parse tree (also referred to as Syntax Trees). However, our method depends only on the dependency graph of the sentence. We use dependency relations to determine the order of tokens that are required to be processed. We then perform some post-processing steps to get the final Text for Indian Sign Language. In this, I also introduce novel works in the field of sign language conversion mentioned below.

1.5 Major Contributions

1. **Larger Standard Data**: Up till now most translations that are done from English to ISL rely on very small datasets typically of up to 1000 words[31][11][8]. We have analysed over 5000

sentences and developed rules which are more robust. This will be a better basis for anyone trying to tackle this problem in the future.

2. **Tackle Complicated Sentences**: Up till now, most research has focused on grammar rules but this method for a large part has been restricted to very simple sentences. [31] which is the latest paper in this uses a rule-based approach to translate English sentences to ISL videos. The rules are all syntactic transformations on the constituency parse tree. We try to tackle more complicated sentences here which are simple in structure but some of the meanings conveyed by them are complicated and the previously used methods fall short in some regards.

3. **Framework for handling Multi-Word Expressions**: We have used approach such as Multi-Word Expression detection to handle hidden meanings within the sentence structure. Consider the following phrases: best friend, every day, and black board. The combined meaning of the words has a different meaning from the individual words. Some of these combinations have signs associated with them which are not simple combinations of the individual signs. Hence, these words needed to be identified and dealt with together as a single unit.

4. **WordNet-based Synonym Substitution**: When we were translating the sentences in our dataset, we came upon the following sentence: *He is an excellent doctor*. We could not find a corresponding word for excellent. When we approached an ISL expert with it, they simply translated it as:

```
English: He is an excellent doctor.
SIGN: HE DOCTOR GOOD
```

   Here we see that since *excellent* is not present in ISL, they use the closest word with the same meaning as a substitute. We have used Synonym Substitution to try to replicate the same.

1.6 Thesis Organization

Chapter 2 provides an introduction to Indian Sign Language. We felt necessary to include this section so that readers can appreciate the complexities involved in applying NLP tools to Indian Sign Language given it's various peculiarities and differences from spoken languages.

Chapter 3 provides a detail theoretical background of Machine Translation. We discuss a bit about rule-based machine translation techniques in detail, thus providing the necessary foundation to go through with the rest of the thesis.

Chapter 4 provides a detailed description of the proposed methodology for the conversion of Text for Indian Sign Language

Chapter 5 Analyzes the performance of discussed approaches for Machine Translation of text to Text for Indian Sign Language using Dependency Trees.

Chapter 6 We summarize the contributions of this thesis and point out some possible directions for future work.

Chapter 2

A Brief Introduction to Indian Sign Language

In this chapter, we provide a brief overview of Indian Sign Language. We are doing this because, unlike most Natural Language Processing Tasks, here were are not converting from one spoken language's written form to another spoken language's written form. There are a lot of peculiarities in sign languages in general they differ a lot from spoken languages as spoken languages can not only convey information but also be expressive. However, in sign languages, the goal is to convey information with as little use of gestures as required as moving your hands to perform a gesture is a slower process than uttering words.

2.1 Sign Languages of the World

Most countries have different sign languages, even among those that use the same spoken/written language. Table 2.1 lists a few countries, the languages spoken by them and the Sign Language used by them:

Country or Sub-continent	Spoken Language	Signed Language	Abbr.
United Kingdom	English	British Sign Language	BSL
United States of America	English	American Sign Language	ASL
Australia	English	Australian Sign Language	Auslan
Spain	Spanish	Lengua de Signos Española	LSE
Mexico	Spanish	Lengua de Señas Mexicana	LSM
India	**Hindi/English**	**Indian Sign Language**	**ISL**
Japan	Japanese	Japanese Sign Language	JSL
Middle East	Arabic	Arabic Sign Language	ArSL
Germany	German	Deutsche Gebärdensprache	DGS
France	French	Langue des signes française	LSF

Table 2.1: Sign Languages of the world

Sign languages vary based on geography in several ways, including grammar, vocabulary, and regional dialects. Here are some ways sign languages can vary:

1. **Grammar**: Just like spoken languages, sign languages have their own grammar rules. For example, American Sign Language (ASL) uses a subject-object-verb word order, while British Sign Language (BSL) uses a subject-verb-object word order. In addition, some sign languages use facial expressions and body language to convey meaning, while others rely more on hand gestures.

2. **Vocabulary**: Sign languages have their own unique vocabulary, which can vary depending on the region. For example, Australian Sign Language (Auslan) uses signs for animals and plants that are native to Australia, while American Sign Language (ASL) has signs for American cultural references such as the White House or Thanksgiving.

3. **Regional dialects**: Just like spoken languages, sign languages can have regional dialects. For example, American Sign Language (ASL) has variations in signs used in different parts of the United States, such as the sign for ”mom” which can vary between regions. Indian Sign Language also has variations between the ones used in Karachi, Delhi [37] as well as the one used in urban vs rural populations [14]

4. **Influence from surrounding spoken languages**: Sign languages can also be influenced by the surrounding spoken language. For example, Japanese Sign Language (JSL) has signs that are similar to the spoken Japanese language, while British Sign Language (BSL) has signs that are influenced by the English language.

These are just a few examples of how sign languages can vary based on geography. It's important to note that each sign language is unique and has its own characteristics that make it a distinct language.

2.2 Indo-Pakistani Sign Language

Indian Sign Language (ISL) is a variant of signed languages used primary in the Indian subcontinent. Indian Sign Language is formally known as Indo-Pakistani Sign Language but has evolved into multiple languages namely Indian Sign Language(ISL), Pakistani Sign Language(PSL), West Bengal Sign Language(WBSL), etc. However most of the conceived differences are arguably the result of political disagreements that are not backed by substantial linguistic evidence.

> that sign language varieties in India, Pakistan, and Nepal are distinct but closely related language varieties that belong to the same language family. (...) Further research is needed to determine if this subfamily includes sign language varieties from other countries and if this sub-family can be grouped with other related subfamilies. [36]

However, Dr. Ulrike Zeshan's research conducted in Karachi and in New Delhi clearly states

> sign language varieties in both cities in fact constitute the same language and have identical grammars, to the extent that practically all observations that were originally based on the

initial data from Karachi apply to the Delhi variety as well. Differences between the two are mainly due to the vocabulary and do not concern any of the grammatical observations.[37]

The language we will adhere to here is the Delhi dialect of Indo-Pakistani Sign Language which is also referred to as Indian Sign Language (ISL).

2.2.1 History of Indian Sign Language

It was during the British colonial period that efforts were made to create a standardized sign language for India. In 1880, the first school for the deaf was established in Mumbai by a British missionary named Andrew Jackson. He introduced British Sign Language (BSL) in the school, which was later modified and adapted to suit the Indian context. The signs were influenced by the local languages, including Hindi, Bengali, and Marathi.

In 1913, the Indian National Association for the Deaf (INAD) was established, which played a significant role in promoting and preserving ISL. The INAD brought together deaf people from different parts of India and created a platform for them to share their experiences and develop a common language. The association published a sign language dictionary in 1953, which included over 2000 signs.

In the 1960s, the Indian government recognized ISL as a language and provided funding for the development of sign language education and research. Several schools for the deaf were established across the country, and sign language was included in the curriculum. The Indian Sign Language Research and Training Center (ISLRTC) was established in 2015, which is a national centre for the research and development of ISL.

2.2.2 Importance of Sign Language for Education

Despite being a robust process, research indicates that the rate at which children acquire language is sensitive to the amount and quality of input they receive from the adults around them. While the quantity of input is important, the quality of input has a greater impact on language development. In most cases, language acquisition occurs naturally as children are exposed to others speaking around them. Research has established a link between certain cognitive skills such as working memory, phonological processing skills, and morphological awareness and the early vocabulary growth of children [29][27]. The development of these cognitive skills, along with environmental factors, plays a vital role in a child's initial vocabulary and vocabulary growth. However, this development is largely dependent on social or academic opportunities and interactions. It is important to note that the standard language acquisition system may not be suitable for all communities, especially those who are dealing with various disabilities. Hence the formal education of Indian Sign Language becomes even more important in these cases.

In [32] Madan M Vasistha conducted a survey that reported that many schools in India were not using standardized ISL signs. The students are taught to make certain gestures. After 20 years, the education of sign languages in these schools started using regional languages and English. These languages are spelt out as the signs for most of the words are not available. In recent years, this has been reported that ISL along with its vocabulary and syntax is standardized by the aforementioned ISLRTC.

A majority of those enrolled in sign language schools, the second language (L2) for most of the Indian deaf community is English (ISL being considered as the first language (L1)) as it is the most prevalent language in the teaching and study materials. The goal of the present research is to take a few steps towards developing an L2 (English) to L1 (ISL). It has been observed that proficiency in L1 actually affects proficiency in L2. [24] [27] [21]. Hence it is critical to develop tools such as the one presented in this paper which facilitates interchange between English and ISL.

2.2.3 Sign Language Typology

The kinds of situations known to hold between spoken languages are not present in sign languages. It is therefore very difficult to apply the traditional historical-comparative method to sign languages. It is impossible for a number of reasons. For example, there is no known equivalent to the regular sound change in sign language. The only way to link sign languages is through their history which has been well recorded due to the fact that most of these languages started out as special education systems often being brought from abroad. For example, American Sign Language is believed to have arisen from a creolization situation involving French Sign Language and pre-existing local sign varieties. [39]. Indian Sign Language's history is well documented and we can hence figure out its typology and relation to other sign languages

2.2.4 Indian Sign Language Grammar

Sign languages resemble the surrounding spoken languages since most sign language users are bilingual to some extent. This influence of the spoken language on sign language, leads to borrowing of from the spoken language. The most common borrowing that we can see is that sign languages mirror the structure of the spoken language. There are systems such as Signed English and Signed Japanese (these are different from British Sign Language and Japanese Sign Language respectively) which use sign language vocabulary, but with the signs appearing in the order of the corresponding sentences from the spoken language. These are however used mostly in educational settings.

2.2.4.1 Morphology

Morphological typology usually recognizes three canonical types of language: isolating, agglutinating, fusional and polysynthetic. These types of traditional typologies cannot be applied to sign languages. This does not mean that sign languages do not have morphological information with the signs. Instead of inflectional morphemes now there are variations in the ways signs are performed and the movement, shape and direction of a sign being performed carry extra meaning along with the word itself which. This phenomenon is not unique to Indian Sign Language only but is seen in other sign languages such as American Sign Language (ASL) too.

Handshape is one of the most important components of ISL signs. There are over 40 different handshapes in ISL, each of which represents a specific meaning or category of objects. For example, the "A" handshape can be used to represent people, while the "S" handshape can be used to represent snakes or thin objects. Handshapes can be combined to create compound signs, such as "spoon" or "television."

Movement is another important component of ISL signs. Movements can be linear or circular and can convey information about the action being performed or the direction of movement. For example, the sign for "to go" might include a forward movement, while the sign for "to come" might include a backward movement.

Location is also an important component of ISL signs. Signs can be made at specific locations on the body or in the signing space, and these locations can convey information about the object being referred to or the location of the action. For example, the sign for "house" might be made by pointing to a specific location in the signing space, while the sign for "shoe" might be made by indicating the location on the foot where the shoe is worn.

Orientation is the final component of ISL signs. Signs can be made with the palm facing up, down, or to the side, and this orientation can convey information about the object being referred to or the action being performed. For example, the sign for "water" might be made with the palm facing down to indicate the fluid nature of water, while the sign for "to eat" might be made with the palm facing up to indicate the act of lifting food to the mouth.

Overall, the morphology of signs in ISL is complex and multifaceted. Signs are created using a combination of handshape, movement, location, and orientation, and these components can be modified to convey different shades of meaning.

ISL also uses inflection to convey meaning. For example, the same sign for "go" can be inflected to indicate past or future tense (discussed in more detail in section 2.1.3.5). The use of facial expressions is also important in ISL, as they can change the meaning of a sign. For example, a neutral face while signing "I'm sorry" would convey a different meaning than a sad or apologetic expression.

2.2.4.2 Fingerspelling

ISL contains 26 signs which can be used to spell out words from the English language. Fingerspelling is an important aspect of Indian Sign Language (ISL) and is used in a variety of ways. Here are some of the common uses of fingerspelling in ISL:

1. Proper nouns: Fingerspelling is often used to spell out proper nouns such as names of people, places, or organizations. For example, the name "John" would be fingerspelled using the appropriate handshapes for each letter of the name.

2. Technical vocabulary: Fingerspelling is also used to spell out technical vocabulary or jargon that may not have a commonly understood sign. For example, a scientific term or a legal term may be fingerspelled to ensure that everyone in the conversation understands the meaning.

3. Emphasis: Fingerspelling can be used to add emphasis to a particular word or concept in a sentence. For example, a signer might fingerspell a word to draw attention to it or to make sure it is understood.

4. Clarification: Fingerspelling can also be used for clarification, especially when the meaning of a sign is unclear or ambiguous. For example, a signer might fingerspell a word to make sure that everyone understands the meaning.

   Teaching: Fingerspelling is an important tool for teaching ISL to beginners, as it provides a visual reference for the letters of the alphabet and helps students to understand the structure of the language.

It is important to note that fingerspelling in ISL follows a different grammar than spoken language. In ISL, fingerspelling is not used as a direct translation of spoken language, but rather as an integral part of the grammar of the language. Fingerspelling in ISL has its own distinct rules and conventions, and it is important for learners to understand these rules in order to use fingerspelling effectively in their communication.

2.2.4.3 Word Reduplication

Word reduplication is a unique phenomenon seen in Indian Languages [1]. It refers to a partial or complete repition of a word or a lexeme. In ISL, we can see complete reduplication of sign which has a combined meaning that is a modified meaning to the original sign. In Indian Sign Language we see reduplication to mark distributive aspect and iterative aspect of a verb.

In ISL, word reduplication can be used to indicate intensity or to show a change in the state or quality of an object or action. For example, the sign for ”eat” can be reduplicated to indicate ”eat a lot,” or ”eat quickly.” Similarly, the sign for ”cold” can be reduplicated to indicate ”very cold.”

Word reduplication can also be used to create new signs by modifying existing signs. For example, the sign for ”flower” can be reduplicated and modified to create the sign for ”bouquet.”

In addition, word reduplication in ISL can be used to indicate repetition or duration of an action or event. For example, the sign for ”to drink” can be reduplicated to indicate drinking continuously, or for a long time. For example:

```
SIGN: I GO MONTH MONTH
English: I go every month
```

Similarly to the distributive aspect, the iterative aspect is also formed by repeating the signs. **However, reduplication is not a phenomenon restricted to Indian Sign Language only** as seen in [10].

We see that similar patterns are present in other languages also. [4] for Swedish Sign Language, [23] for German Sign Language show that not only is reduplication a universal phenomenon in sign languages but also the intention of their use is the same as well.

2.2.4.4 Sentence Structure

Indian Sign Language is supposed to have a free word order similar to Indian languages such as Hindi. However, the sentences mostly follow an SOV(Subject Object Verb) structure mostly in the case of sentences with a single predicate.

```
SIGN: I DEAF
English: I am deaf
```

```
SIGN: I APPLE EAT
English: I eat apple
```

American Sign Language (ASL) on the other hand follows SVO which is the sentence structure used in English. However, there is no known correlation between sign languages and their corresponding spoken languages. We can take the example of British Sign Language (BSL) which has a different sentence structure to ASL and English.

2.2.4.5 Temporal Relations

Temporal relations in Indian Sign Language (ISL) are typically handled through the use of adverbs, non-manual markers, and inflection.

Adverbs are used in ISL to indicate the time frame of an action or event. For example, the adverb ”yesterday” might be signed by pointing backwards over the shoulder, while the adverb ”tomorrow” might be signed by pointing forward. Other adverbs, such as ”now,” ”soon,” and ”later,” can also be used to indicate temporal relations.

Non-manual markers are another important aspect of temporal relations in ISL. These markers include facial expressions and body language that convey information about the timing and duration of actions. For example, a raised eyebrow might indicate a question about when an action will occur, while a nod of the head might indicate agreement about a future event.

Inflection is also used in ISL to indicate temporal relations. Signs can be inflected to indicate past, present, or future tense, and to show duration or frequency. For example, the sign for ”to run” might be inflected to indicate that the action happened in the past, or that it will happen in the future.

In addition to these strategies, temporal relations in ISL can also be conveyed through the use of classifiers. Classifiers are handshapes that represent a category of objects, such as vehicles or animals, and can be used to show the location, movement, and shape of objects in the signing space. Classifiers can be used to show the passage of time, such as the movement of the sun or the rotation of the Earth.

Overall, the handling of temporal relations in ISL is multifaceted and relies on a combination of adverbs, non-manual markers, inflection, and classifiers to convey meaning about the timing and duration of actions and events.

2.2.4.6 Relative Clauses

Relative clauses are subordinate clauses that contain some extra information about the main clause. They are shown in sign language similar to adjectives and adverbs i.e. they follow the phrases they subjunct. In ISL, relative clauses are formed by positioning the clause after the noun that it modifies.

```
SIGN: MAN SHIRT BLUE WEAR
English: the man who is wearing a blue shirt
```

Another strategy is to use non-manual markers such as eye gaze, head nodding, and eyebrow-raising to indicate relative clause relationships. [28] also notes that ISL has a tendency to omit relative pronouns, instead relying on context and spatial referencing to convey meaning.

Overall, the handling of relative clauses in ISL is complex and relies on a combination of spatial referencing, classifiers, and inflection to convey the intended meaning. Successful communication of relative clauses in ISL requires a strong understanding of these linguistic tools and how they can be used to convey information in a clear and effective way.

2.2.4.7 Interrogative sentences and Wh Questions

ISL has only one sign for marking interrogation. We refer to it as the general wh sign (G-WH) and it is placed along with common noun signs to denote what kind of answer the question expects. This leads to the formation of signs such as:

- $Where \rightarrow PLACE + G - WH$
- $When \rightarrow TIME + G - WH$
- $Who \rightarrow PERSON + G - WH$

Wh questions follow the same structure as normal sentences but the question word is put at the end of the sentence. [28] [38]

2.2.4.8 Complex Sentences

A complex sentence is a sentence that consists of one independent clause and at least one dependent clause. An independent clause is a clause that can stand alone as a complete sentence, while a dependent clause relies on the independent clause to complete its meaning.

In a complex sentence, the dependent clause typically functions as a modifier or as a subordinate element to the independent clause. The dependent clause can provide additional information, clarify the meaning of the independent clause, or show the relationship between two ideas. ISL tries to handle complex sentences by using non-manual features to express the simultaneous nature of the actions as ISL does not contain words like *if, while and though*. However, some experts we consulted, from different parts of the country, when asked to translate sentences, independently chose to fingerspell these words such as if *if* which seem to indicate that it has evolved into an accepted way to convey those ideas. It is to be noted here that the word *if* itself has no candidate in the video dictionary.

Overall, the grammar of ISL is complex and nuanced, just like any spoken language. It has its own rules for sentence structure, word order, and inflection, and relies on non-manual markers, facial expressions, and classifiers to convey meaning.

Chapter 3

Approaches to Machine Translation

3.1 Machine Translation

The idea of machine translation (MT) started in the seventeenth century, but it became possible only in the middle of the twentieth century when there was a growth in the computing capabilities of machines [13]. The early systems were developed using the three basic approaches of MT: the direct approach, the rule-based transfer approach and the Interlingua approach. All these approaches require deep linguistic analysis, transformation rules written by language experts and various tools such as parsers, morphological analyzers, large bilingual dictionaries etc. to facilitate the translation process.

3.1.1 Statistical Machine Translation

Statistical machine translation, or SMT, is a data-driven approach to machine translation that involves training models on large parallel corpora, which are collections of texts in the source and target languages. These models use statistical algorithms to identify patterns and relationships between the source and target language data and then generate translations based on these patterns.

SMT has several advantages, such as its ability to handle different language pairs and its scalability to handle large amounts of data. However, it also has some limitations, such as the requirement for large amounts of high-quality training data and the potential for errors due to the use of statistical models.

3.1.2 Neural Machine Translation

Neural machine translation, or NMT, is a type of machine translation that uses artificial neural networks to learn the mapping between the source and target languages. NMT models are typically trained on large parallel corpora and use multiple layers of neural networks to encode the source language input and decode the target language output.

NMT has several advantages, such as its ability to produce fluent and natural-sounding translations and its ability to handle a wide range of language pairs. However, it also has some limitations, such as the requirement for large amounts of training data and the potential for errors due to the complexity of the models.

3.1.3 Hybrid Machine Translation

Hybrid machine translation, or HMT, is an approach to machine translation that combines the strengths of multiple approaches, such as rule-based, statistical, and neural machine translation. HMT models use a combination of these approaches to produce translations that are more accurate and natural-sounding than those produced by individual approaches.

HMT has several advantages, such as its ability to handle a wide range of language pairs and its ability to produce high-quality translations. However, it also has some limitations, such as the complexity of the models and the difficulty in combining the different approaches.

3.1.4 Rule Based Machine Translation

Rule-based machine translation (RBMT) is the oldest approach to machine translation, which involves using a set of rules to analyze the structure and meaning of the input text and then generate the corresponding output text in the target language. These rules are created by human experts who have a deep understanding of both the source and target languages. [5] [15] [34] are some examples of rule-based systems in Sign Language Translation.

The rule sets for RBMT are typically created by linguists who break down sentences into their grammatical components, such as nouns, verbs, adjectives, and prepositions, and create rules that govern how these components can be combined to form meaningful sentences in the target language. These rules are stored in a database and used by the RBMT system to generate translations.

One of the key advantages of RBMT is its ability to handle complex sentence structures and produce grammatically correct translations. RBMT can also be useful when dealing with specialized domains, such as legal or medical texts, where precise and accurate translations are critical.

However, RBMT also has some limitations. One of the main limitations is the difficulty in creating and maintaining the rule sets for all possible language pairs. This is because each language has its own unique grammar rules and syntax, and it can be challenging to create and maintain rule sets that are comprehensive and accurate. As a result, RBMT systems can be expensive and time-consuming to develop and maintain.

Another limitation of RBMT is that it does not take into account the context and cultural nuances of the source language. This can lead to translations that are technically correct but may not accurately convey the intended meaning or tone of the original text.

Despite these limitations, RBMT can still be a useful approach for certain translation tasks, especially when dealing with structured and technical content.

The process of RBMT translation follows the principles of Analysis-Transfer-Generation (ATG). The ATG mechanism is proposed by Benard Vauquois through a triangle known as the Vauquois trian-

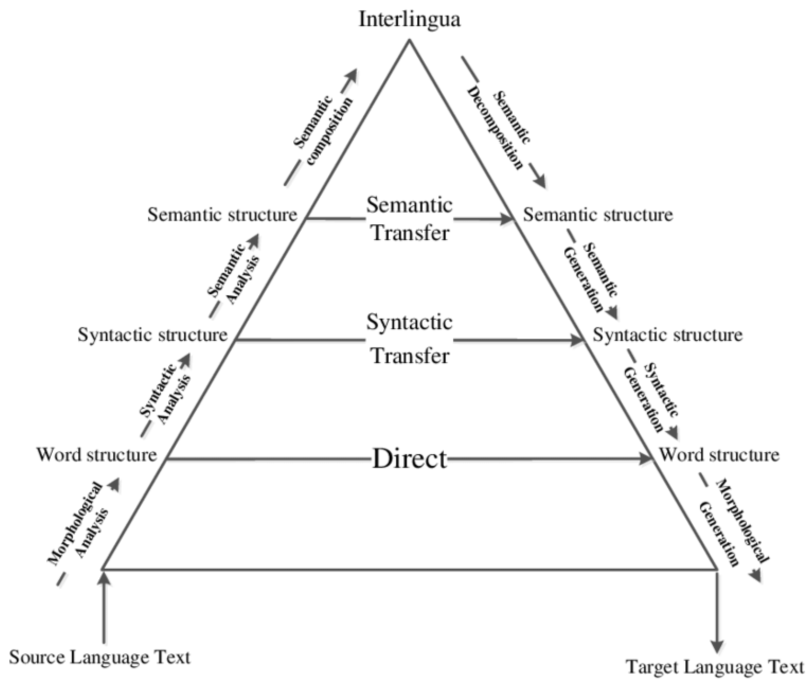


Figure 3.1: The Vauquois Triangle

gle [33]. The Vauquois triangle shows how translation can take place between two languages at different levels of linguistic analysis. For example, a translator might start with a source text in one language, analyze its syntax, and then translate it into the syntax of the target language. Alternatively, a translator might start with the meaning of the source text, analyze its semantics, and then translate it into the semantics of the target language. The Vauquois triangle has been used to develop and analyze various approaches to machine translation and natural language processing. It highlights the complex relationships between language levels and the challenges involved in accurately translating between them.

There are three main ways of doing rule-based machine translation:

1. Direct Translation
2. Interlingua based Translation
3. Transfer based Translation

3.1.4.1 Direct Translation

In direct translation, we proceed word-by-word through the source language text, translating each word as we go. The direct translation uses a large bilingual dictionary, each of whose entries is a small program with the job of translating one word. In transfer approaches, we first parse the input text and then apply rules to transform the source language parse structure into a target language parse structure.

We then generate the target language sentence from the parse structure. In the interlingua approach, we analyze the source language text into some abstract meaning representation, called an interlingua. We then generate the target language from this interlingual representation.

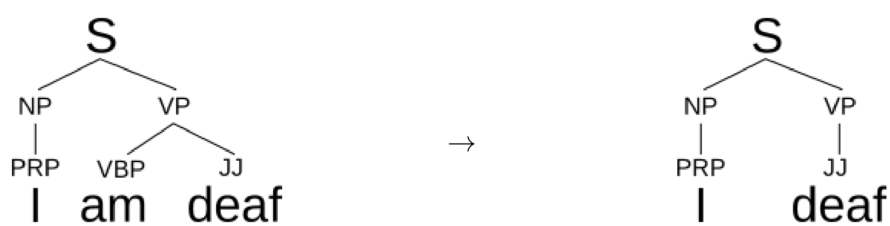
3.1.4.2 Interlingua based Translation

Interlingua-based translation is a type of machine translation that uses the Interlingua language as an intermediate representation for translating between multiple languages. The Interlingua language is designed to represent the meaning of a sentence in a language-independent way, which allows it to serve as a common intermediary between different languages.

The basic idea behind Interlingua-based translation is to first analyze the source language sentence to generate an Interlingua representation of its meaning, and then use that representation to generate a target language sentence. This approach allows for more accurate translations because the meaning of the source sentence is captured in the Interlingua representation, rather than just the words themselves.

3.1.4.3 Transfer based Translation

Transfer-based translation is a type of machine translation that involves transferring linguistic structures or rules from the source language to the target language. In transfer-based translation, the source language sentence is first analyzed to identify the linguistic structures, such as grammatical rules, syntactic structures, and semantic representations. These structures are then transferred to the target language, where they are used to generate a translation.



(a) Phrase structure tree in English (b) Phrase structure tree in ISL

Figure 3.2: Phrase structure tree in English and ISL

In 3.2a we see the phrase structure tree constructed for the simple sentence *I am deaf* in English. We see that each sentence should have a noun phrase (nominally subject here) and a verb phrase. The noun phrase here is

3.2 Literature Review

As we have alluded to before, an automated spoken language to sign language translation system can help promote accessibility, inclusivity, efficiency, and accuracy in communication, particularly for those who rely on sign language as their primary means of communication. As such automated spoken language to sign language translation systems have been in the works for a long time, especially for sign languages such as ASL and BSL and it is worth looking at how these systems work and look at the positives and negatives of these systems and see how we can improve upon these

3.2.1 The Challenges of Cross-Modal Translation: English-to-Sign-Language Translation in the Zardoz System

The Zardoz system [34], developed by MIT, was designed for English to Japanese Sign Language (JSL), American Sign Language (ASL) and Irish Sign Language in the domain of health. One challenge with such a system is the variability of sign languages. Sign languages vary widely between different countries and regions, and even between different communities within the same country. This means that any system for English-to-generic-sign-language translation would need to be able to adapt to different sign-language dialects and variations. Another challenge is the lack of standardization in sign language notation. While written languages have standardized systems of spelling and grammar, sign languages do not have a universal notation system. This makes it difficult to develop automated translation systems that can be used across different sign language communities.

The proposed architecture is shown in Fig 3.3. The system tries to implement an interlingua-based system. The input sentence is first passed through what we would call a morphological analyzer which they call Lexical Experts. We have to assume that there was some sort of tokenization step before this which is not specified in the paper. After the morphological analyzer, the tokens are passed through an Idiomatic Preprocessor where known idioms are replaced with equivalent phrases which are better suited for parsing. The tokens are then used to generate a syntactic parse tree which is then used along with the morphological features to build some sort of a dependency relation graph of the tokens. These are then put into a schematic representation of the sentence which is the basis of the next set of mappings to transform the sentence into the list of signs which need to be fed to the animator. For the actual animations, the system uses a Doll Control Language (DCL) which is used to animate a virtual avatar for animations. From our understanding, this seems to be some sort of precursor to SiGML [41] and HamNoSys [12]. The Zardoz project did not provide any concrete experimental results.

3.2.2 Prototype Machine Translation System From Text-To-Indian Sign Language

The authors [8] have created a translation system that converts English text to Indian sign language using transfer grammar rules. The whole system consists of 2 subsystems: ISL MT Architecture and SL Dictionary.

The MT system is based on machine translation that uses four main modules: input text preprocessor and parser, LFG f-structure representation, Transfer Grammar Rules for ISL sentence generation and ISL synthesis. The entire architecture is described in Fig

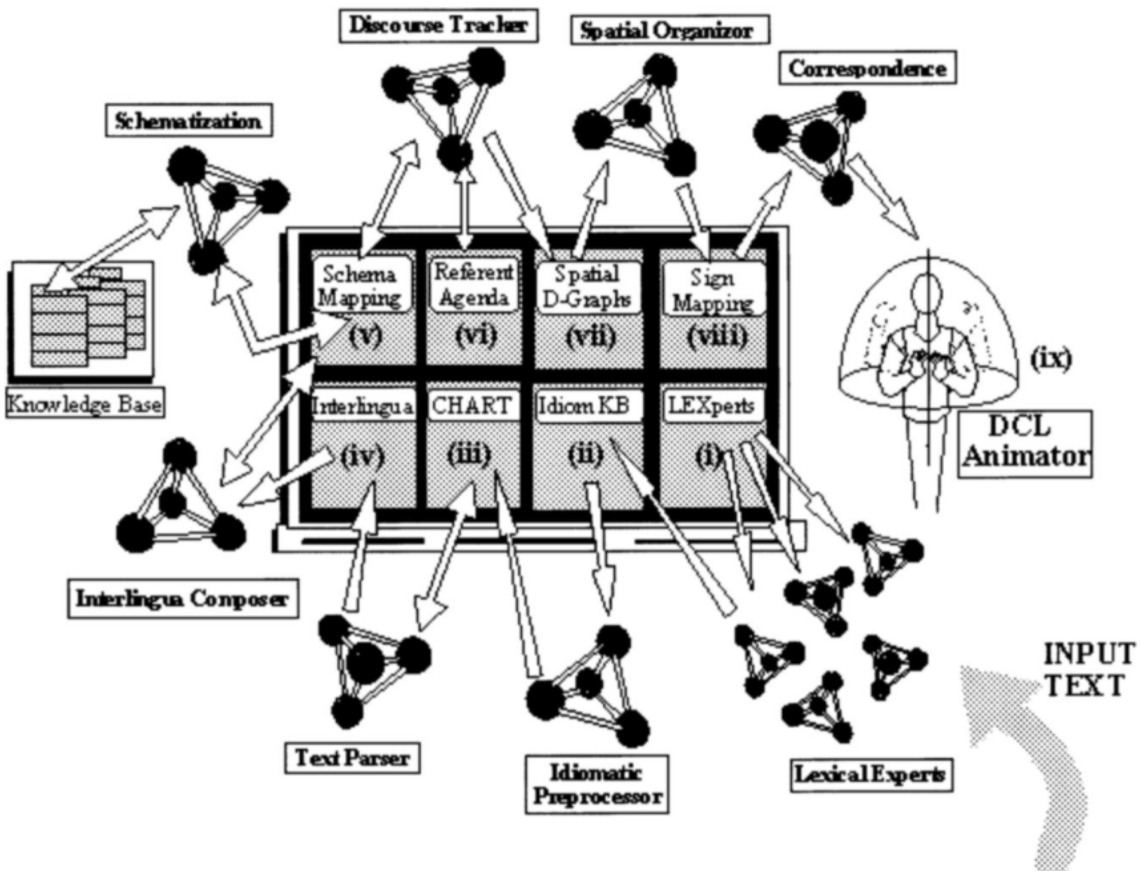


Figure 3.3: Zardoz Architecture

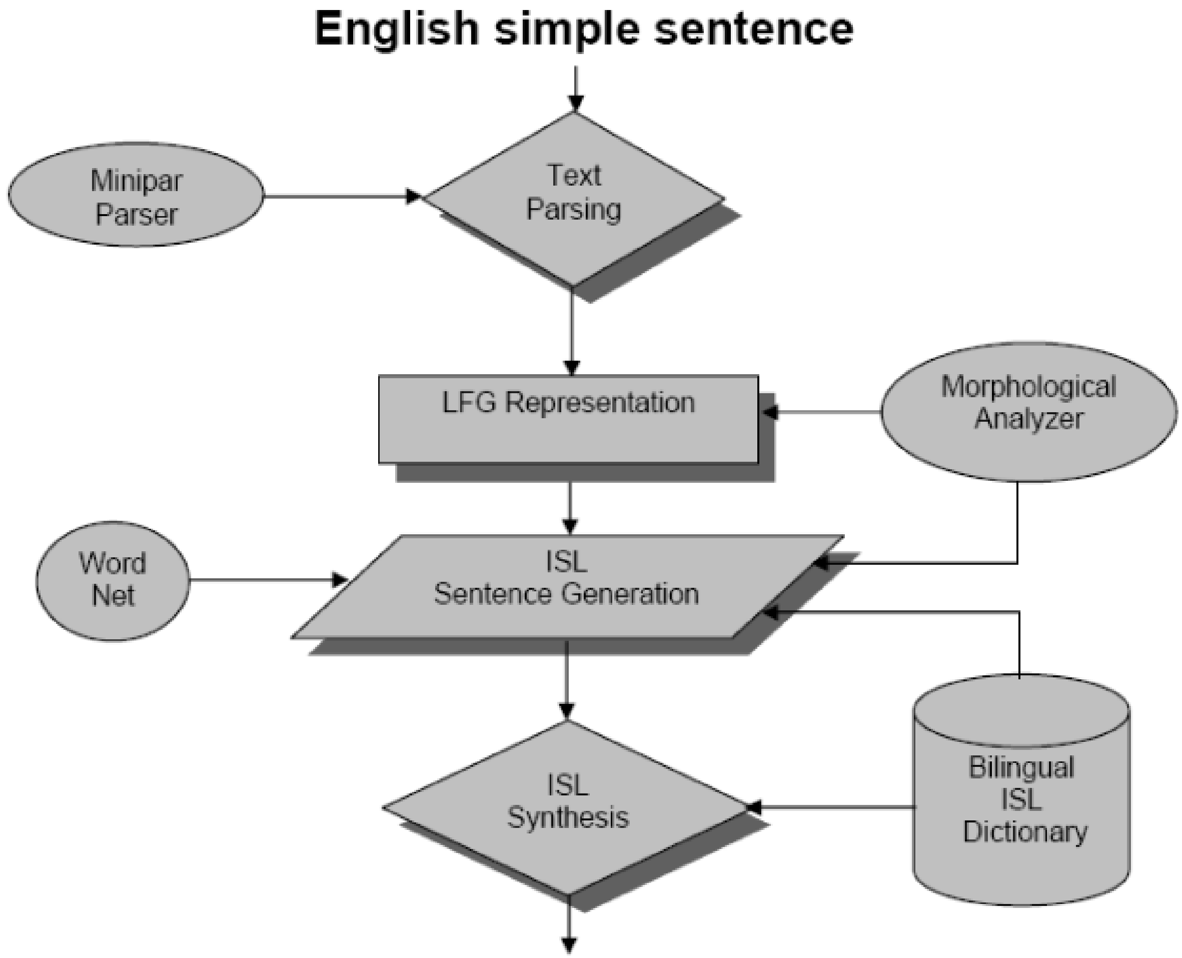


Figure 3.4: Architecture of the Text-to-ISL MT system in [8]

The preprocessing stage identifies the frozen phrases and temporal expressions, before sending them to the Minipar parser. A phrase lookup table with approximately 350 phrases is used for this. The Minipar parser is used for dependency parsing to generate the LFG f-structure of English. The English f-structure is converted to ISL f-structure by applying proper transfer grammar rules. The rules are not mentioned in the paper. In the ISL Sentence Generation stage, the sentence goes through a lexical selection and word order correspondence. Lexical selection is done using English-ISL bilingual dictionary. For example, a word like *DINNER* in English is replaced by *NIGHT FOOD* in ISL and *MUMBAI* is replaced by the sign of *BOMBAY*.

The system claims an accuracy of 96.37% for sentences without directional verbs and 92.53% for compound sentences. The overall accuracy is 89.4% over 208 sentences.

The goal of the SL Dictionary is to Build a cross-platform multilingual multimedia SL-Dictionary tool that can be used to create a large SL lexicon. The architecture of the SL dictionary tool has been illustrated in Figure 3.5. The system has two modules: Expert module and User Module. The input to the system may be a word, phrase, or sentence. If the input is a word, the system identifies all possible semantic senses of the word using a wordnet, along with the part of speech (POS) of that word. The

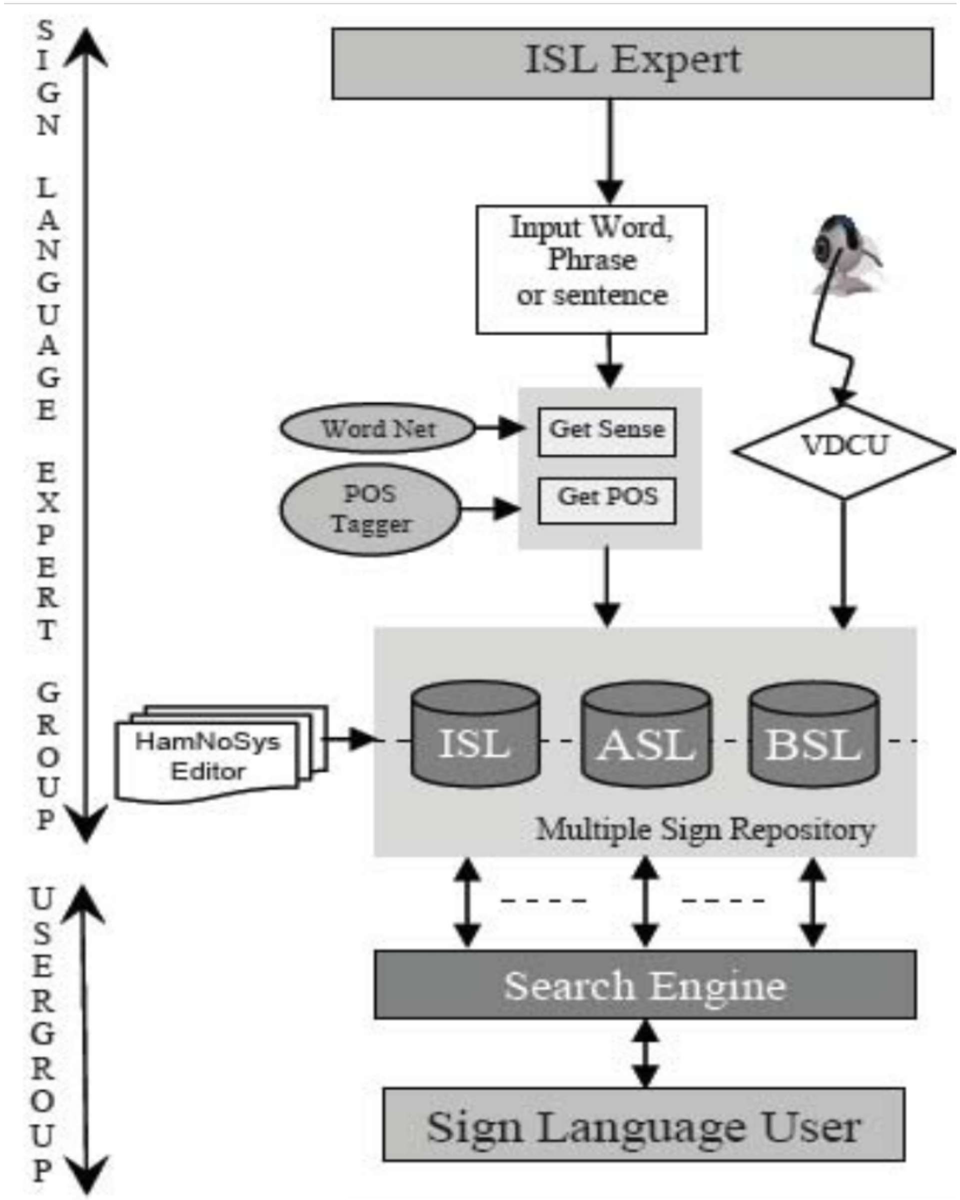


Figure 3.5: SL Dictionary Architecture in [8]

issue with the dictionary is that as we have mentioned before, sign languages in different parts of the world are not really compatible, so the point of mixing up the signs from different languages is kind of pointless.

3.2.3 Automatic Translation of English Text to Indian Sign Language Synthetic Animations

The paper [11] presents the prototype for English Text to Indian Sign Language conversion system using synthetic animations in the real domain.

The overall architecture of the system is shown in the figure 3.6. It consists of 7 modules:

1. English parser for parsing the English text. This step generates a syntax tree from the sentence using the Stanford Parser.

2. Sentence reordering module based on ISL grammar rules. The grammar transfer rules used by the system are in table 3.1:

Verb Pattern	Rule	Input Sentence	Parsed Sentence	Output Sentence
verb + object	VP NP	go school	(VP (VB Go) (NP (NN school)))	school go
subject + verb	NP V	birds fly	(NP (NNS birds)) (VP (VBP fly))	birds fly
subject + verb + subject complement	NP V NP	his brother became a solider	(NP (PRP his) (NN brother)) (VP (VBD became) (NP (DT a) (NN soldier)))	his brother a solider became
subject + verb + indirect object + direct object	NP V NP NP	I lent her my pen	(NP (FW i)) (VP (VBD lent) (NP (PRP her)) (NP (PRP my) (NN pen)))	i her my pen lent

Table 3.1: Examples of Grammatical Reordering of Words of English Sentence in [11]

3. Eliminator for eliminating the unwanted words. Indian sign language sentences are formed of content words. All the functional words like linking verbs, suffixes, and articles are not used and hence are removed after the grammar transfer.

4. Lemmatization for getting the root word of each word since ISL uses the root words in their sentences. All the words used must not contain suffixes, gerunds, or inflections.

5. Synonym replacement module to replace the unknown word with its synonym counterpart.

6. Word to SiGML conversion using HamNoSys. For each word, it's HamNoSys representation is looked up in the dictionary. Then the HamNoSys codes are converted to SiGML to be used as input for the SiGML player.

7. Synthetic Animation module generates animations from the SiGML tags.

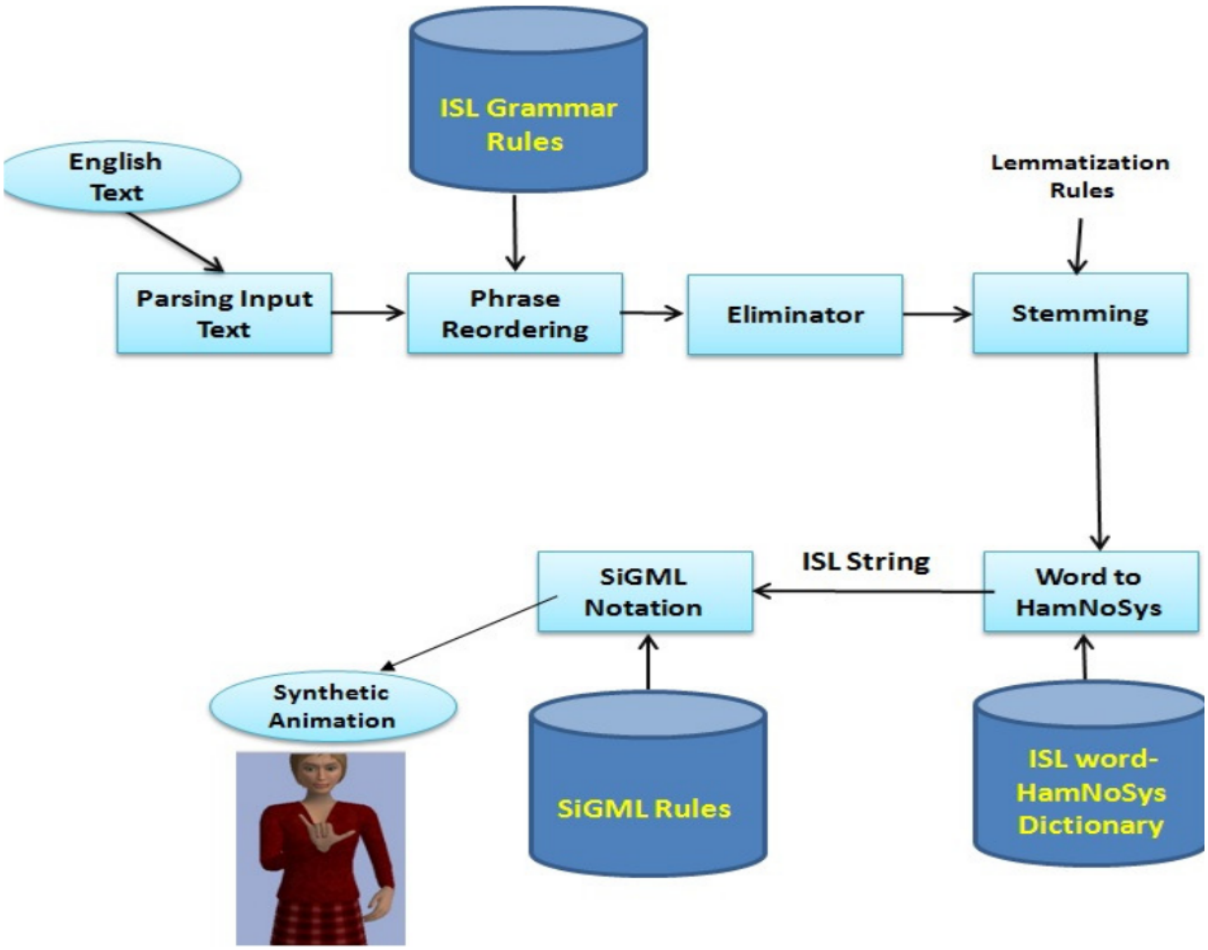


Figure 3.6: English Text to ISL Synthetic Animation System by [11]

The author evaluates the system using a dataset of English sentences and their corresponding ISL translations. The results show that the system is able to generate accurate and natural-looking animations that closely match the original ISL translations. The author also notes that the system has the potential to be extended to other sign languages and can be integrated into existing assistive technologies to provide better communication access for deaf people. The issue with the paper is that the translations are not very accurate as the translation process fails to capture any sort of meaning of the words. The system also cannot handle negative, interrogative and complex sentences.

3.2.4 Sign Language Generation System Based on Indian Sign Language Grammar

[31] is the latest paper on this topic published in 2020. The flow chart 3.7 demonstrates the process by which the system operates. Initially, the user inputs an English word or sentence. The input is then parsed and filtered, resulting in a list of root words. The system then cross-references the morphological information of the root words with pre-existing ISL grammar rules stored in the database. If a named en-

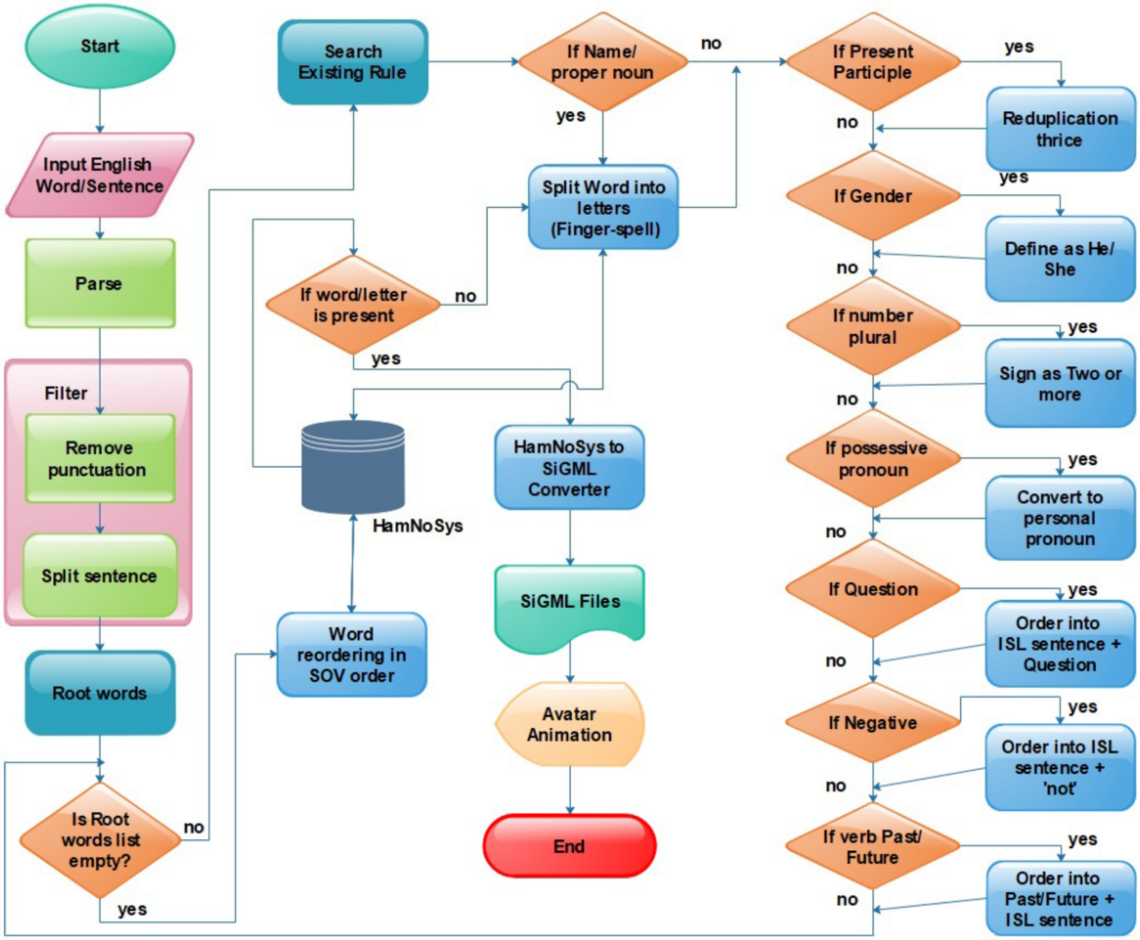


Figure 3.7: Flow chart of the rules in [31]

tity appears or the word is not present in the dictionary, the system finger-spells it. To indicate a present participle, the sign is repeated three times. In a similar fashion, possessive pronouns are converted to personal pronouns, and plural nouns are represented by repeating the sign two or more times. Spatial timelines are used to represent tense in ISL, while he or she is utilized to denote gender. Next, the root words are arranged in SOV order. If a rule for a specific input cannot be found, the system retrieves the SiGML files for the root words and generates ISL signs through avatar animation.

The proposed system has achieved a BLEU score of 0.95 for a corpus of 1,000 sentences. However, there has been no description of the dataset given, where it was sourced from, or the word distribution in the database. We also see that the only rule applied for changing word order is changing a sentence from SVO to SOV but other changes such as reordering of adjectives or adverbs and handling complex sentences have not been addressed. However being the latest one, it was the benchmark for us to improve upon.

3.3 Our Approach

As we have seen, most methods actually rely on syntax trees of the sentence to generate meaning but we are using dependency trees to get the relation between words in the sentence and then we place how the words are ordered relative to the other words. For example, since we know that ISL follows the SOV format, the root verb must be preceded by the subject and object and then the verb. Another key factor is that the dependency tree of a sentence does not depend on the order in which the words appear in the original sentence. Hence the transformation rules if written for one language will be able to handle all languages for which dependency trees can be constructed. This approach is inspired by the Zardoz system which was not used for ISL and our system is much simpler thanks to those who have developed the tools over the 20 years between that system and ours.

Chapter 4

System Development

4.1 Novelty of our model

As discussed above, most rule-based methods look at syntactic transfer rules to translate sentences from one language to another but we have chosen to use dependency trees for extending these to other languages. Another important addition to our model is the ability to handle named entities as well as co-reference resolution to link pronouns to entities. This greatly improves the understandability of our model. It is to be noted here that these two specific modules also need to have been developed for the source language for the translation to take place. First is the MWE model which detects frozen phrases in a sentence and then instead of translating it word by word, looks for the direct phrase in the bilingual dictionary or the closest word in the wordnet to that word in terms of meaning. Another model which is novel to our system is the use of neural coreference resolution which has been used to identify whether a pronoun used in the sentence is referring to the subject itself or not as such words are replaced by the $self/OWN$ sign in ISL.

4.2 Dataset

We use the SimpleWiki Dataset [1] described in [7] for our research purposes. The dataset was used for training models to translate complex sentences into simple sentences. We are using the corpus of simple sentences from this as rule-based translation systems tend to work best on simple sentences. The data contains a total of 166410 sentences. The data was not clean and contained many incomplete sentences. Hence we filtered out a lot of these sentences. For simplicity purposes, we wrote an automated script to detect sentences that had at least a verb and a subject or an object after which we were left with 30767 sentences. From those, we randomly selected 5592 sentences which were around 5-8 words long.

[1] https://www.kaggle.com/datasets/mfekadu/sentences?select=simple-wiki-unique-no-end-punct-sentences.csv

4.3 System Architecture

We have broken down the entire translation process into three phases:

4.3.1 Pre-Processing

Before we can do transformations on the sentence, we first do some pre-processing to get the required information for the transformations. We are mostly using Stanza [26] which is a collection of accurate and efficient tools for the linguistic analysis of many human languages.

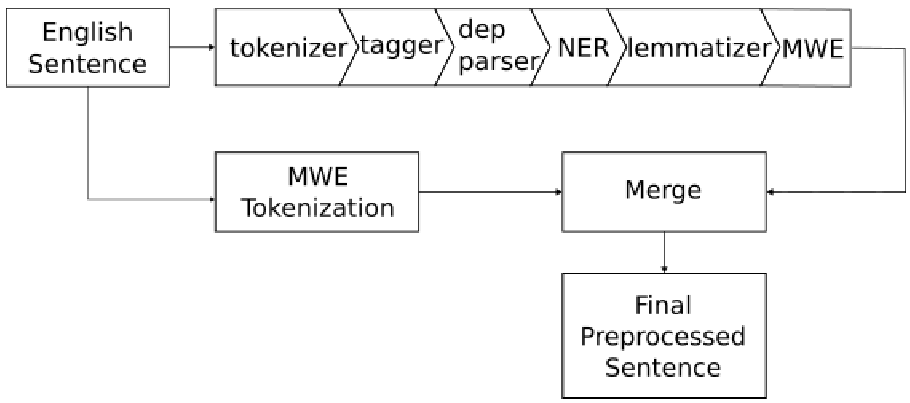


Figure 4.1: Pre-processing Pipeline

4.3.1.1 Tokenization

A tokenizer divides the text into a sequence of tokens, which roughly correspond to ”words”. Stanza combines tokenization and sentence segmentation from raw text into a single module. This is modelled as a tagging problem over character sequences, where the model predicts whether a given character is the end of a token, the end of a sentence, or the end of a multi-word token.

4.3.1.2 POS Tagging and Morphological Features Analyser

A Part-Of-Speech Tagger (POS Tagger) is a piece of software that reads the text in some language and assigns parts of speech to each word (and other tokens), such as nouns, verbs, adjectives, etc., although generally computational applications use more fine-grained POS tags like ’noun-plural’. The English taggers use the Penn Treebank tag set [19]. For each word in a sentence, Stanza assigns it a part of speech (POS) and analyzes its universal morphological features (UFeats, e.g., singular/plural,1st/2nd/3rd person, etc.). To predict POS and UFeats, they adopt a bidirectional long short-term memory network (Bi-LSTM) as the basic architecture.

4.3.1.3 Dependency Parsing

A dependency parser analyzes the grammatical structure of a sentence, establishing relationships between ”head” words and words that modify those heads. Stanza parses each sentence for its syntactic structure, where each word in the sentence is assigned a syntactic head that is either another word in the sentence or in the case of the root word, an artificial root symbol. It implements a Bi-LSTM-based deep biaffine neural dependency parser [9]. It further augments this model with two linguistically motivated features: one that predicts the linearization order of two words in a given language, and the other that predicts the typical distance in linear order between them. These features significantly improve parsing accuracy [25].

4.3.1.4 Named Entity Recognition

Recognizes named entities (person and company names, etc.) in text. This annotator primarily employs one or multiple machine learning sequence models to assign labels to entities, although it may also utilize specialized rule-based components, particularly for identifying and interpreting dates and times. The entities that can be identified depend on the language being used, and the number of recognized entities is typically more restricted for languages other than English. Usually, the annotator runs several named entity recognizers and then combines their outcomes, but it can also run only one annotator or a rule-based quantity NER. By default, the annotator recognizes 12 classes of entities for English, which include named entities such as PERSON, LOCATION, ORGANIZATION, and MISC; numerical entities such as MONEY, NUMBER, ORDINAL, and PERCENT; and temporal entities such as DATE, TIME, DURATION, and SET. For NER we adopt the contextualized string representation-based sequence tagger from [3].

4.3.1.5 Lemmatization

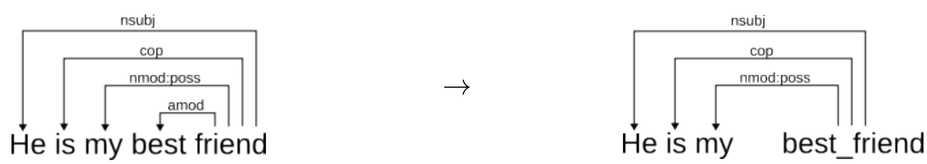
Stanza also lemmatizes each word in a sentence to recover its canonical form (e.g., did $\rightarrow$ do). Similar to the multi-word token expander, Stanza's lemmatizer is implemented as an ensemble of a dictionary-based lemmatizer and a neural seq2seq lemmatizer. An additional classifier is built on the encoder output of the seq2seq model, to predict shortcuts such as lowercasing and identity copy for robustness on long input sequences such as URLs.

4.3.1.6 Multi-word expression recognition

I go to school every day $\rightarrow$ I school go every day Multi-word expressions are groups of words that denote a single idea. Detecting MWEs is important in ISL because their order does not change during translation. Every day would become day every according to the rules but it should not. To detect this, we used jMWE's Standard MWE Index data file which is a library for detecting Multi-Word Expressions (MWE) in text described in the paper [17]. It contains 67037 MWEs collected from various online sources along with the POS tag for the entire unit. These tokenizations happen in parallel to the tokenization above using NLTK MWETokenizer as described in [18]

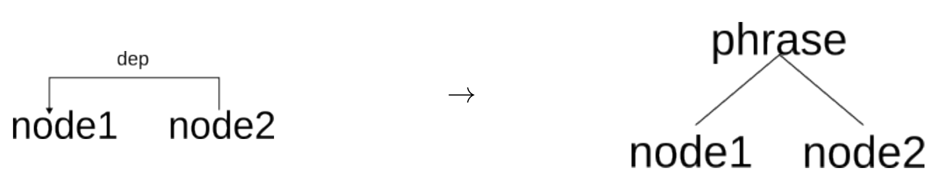
4.3.1.7 Merging the Two Pipelines

At the end of the two pipelines, we need to merge the MWE tokenization information into the information received from Stanza. So the problem is how to define the dependency relation of the MWE token with respect to the rest of the sentence. So for this, we take the tokens from the dependency tree and select the head of the subtree as the representative of the group. We preserve the dependency of the head and remove the dependencies and POS information of the rest of the tokens as they will henceforth be processed as a single token. This process is somewhat inspired by [16].

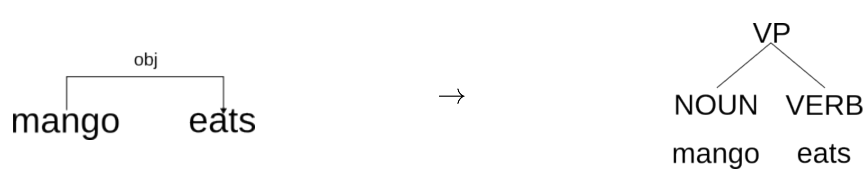


4.3.2 Grammar Transfer Rules

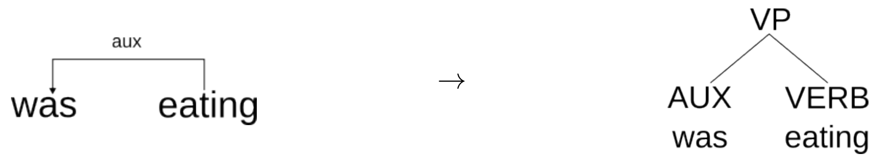
Here we are trying to convert dependency relations to syntactic relations. The dependency relations in the source language are defined as:
$node_1(dependency) \rightarrow node_2$ which are then converted to syntax tree fragments with the notation $phrase\ tree\ branch \rightarrow (node_1)(node_2)$



1. Source: $VERB(root) \rightarrow NOUN(obj)$
   Transformation: $VP \rightarrow [NOUN(obj)][VERB(root)]$

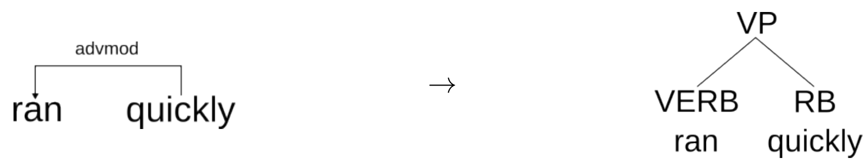


   Explanation: Since ISL follows SOV structure, the sentence arrangement should be NP NP VP. That is why we put all the nouns to the left of the verb.



2. Source: $VERB(any) \rightarrow AUX(aux)$
 Transformation: $VP \rightarrow [AUX(aux)][VERB(any)]$
 All auxiliaries to a verb come before the verb

3. Source: $VERB(any) \rightarrow ADV(advmod)$
 Transformation: $VP \rightarrow [VERB(any)][ADV(advmod)]$

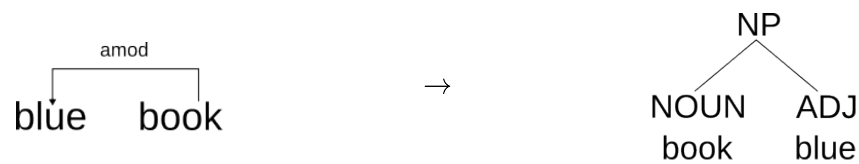


All adverbs to a verb come after the verb

Note: Rules 2 and 3 would not have appeared in syntactic transformation rules because they do not change the order of the syntax tree. But since we want to extend this to other languages, we need to specify where each child node goes wrt to the head node.

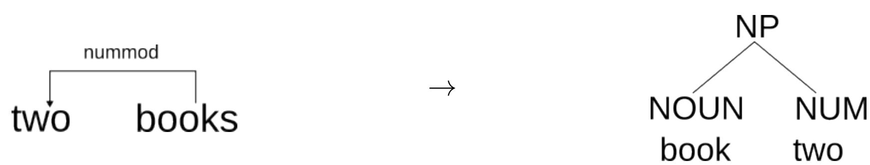
4. Source: $VERB(root) \rightarrow VERB(any)$
   Transformation: $VP \rightarrow [VERB(root)][VERB(any)]$
   This case is written to handle multiple clauses in a sentence. In case of multiple clauses in a sentence, the clauses will be processed in the order.

5. Source: $NOUN(any) \rightarrow ADJ(amod)$
   Transformation: $NP \rightarrow [NOUN(any)][ADJ(amod)]$



   Adjectives follow the noun they describe

6. Source: $NOUN(any) \rightarrow NUM(nummod)$
   Transformation: $NP \rightarrow [NOUN(any)][NUM(nummod)]$



   Numbers are handled in two ways in ISL:

   (a) Reduplication: Repeat the noun to signal plurality

   (b) Numbers follow the noun denoting the quantity

7. Source: $NOUN(any) \rightarrow ANY(acl : relcl)$
   Transformation: $NP \rightarrow [NOUN(any)][ANY(acl : relcl)]$
   A relative clause is similar to an adjective and hence it comes after a noun.

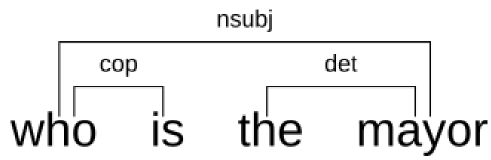
4.3.3 Post Processing

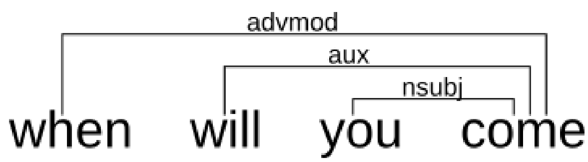
After applying all these transfers we see that the system does produce mostly accurate results for most cases but it still does not work for interrogative sentences and negative sentences. So we make some adjustments required for those cases and then do three additional steps to get the final result.

4.3.3.1 Interrogative and Negative Sentences

Interrogative sentences are sentences that ask a question. In ISL, as discussed previously, interrogative sentences are signed by first showing the signs of the sentence converted to the imperative form

followed by the question word. These questions words usually donot have a specific dependency role assigned to them and can occur in many relations depending on the type of sentences. Examples:
Apart from this, it also becomes very difficult to judge whether or not a wh-word used in a sentence





is a question or not e.g. $I\ know\ what\ he\ is\ talking\ about \rightarrow I\ he\ talk\ about\ know$ [2]. In this sentence, the word *what* is a wh-word but is used as a conjunction. Similarly, questions may sometimes be formed in English without using a wh-word altogether e.g. $Did\ you\ do\ your\ homework \rightarrow you\ homework\ do\ what$. It is difficult to handle such cases and what we have tried to do is just take the wh-word and pick it and place it at the end of the sentence.

Negative sentences are sentences that have a not or no in them. They convey the meaning that the opposite of the event occurs. We apply a similar strategy for negative sentences too by just picking and placing the negative word at the end of the sentence e.g. $He\ is\ not\ a\ doctor \rightarrow he\ doctor\ not$. Similar strategies would work well with other languages too.

As we see, negative sentences and questions are handled similarly in ISL. But then the question arises that if in case we have a negative question sentence such as *Who will not come with me?*, how is it signed in ISL? We did not find any literature discussing this. So we contacted ISLRTC and their group of expert signers seemed to sign the negative word before the question word. So the previously mentioned example would be translated as $Who\ will\ not\ go\ with\ me? \rightarrow me\ with\ go\ not\ who$

4.3.3.2 Synonym Substitution

As we have mentioned earlier, Indian Sign Language is still new and there are many terms that are yet to be added to the dictionary. As such till now there are only 10000 words in the ISL dictionary. As a comparison, English has around 200,000 words [6] which is 20 times the size. Hence there are a lot of words in English which do not have an equivalent sign. If such a word occurs in a sentence, it needs to be translated somehow. For that, we substitute those words that are not in the video dictionary with words close to the original word's meaning and have a corresponding word in the dictionary. The closest words are selected from the synset in the WordNet-based on a score denoting how similar two-word senses are, based on the shortest path that connects the senses in the is-a (hypernym/hyponym) taxonomy. The WordNet we used was the one described by [20] and we use NLTK [18] to access it.

4.3.3.3 Stop-word Removal and Lemmatization

After all the above processing is done, we need to clean up the sentences. As discussed in the previous chapter, ISL does not contain words such as articles and certain functional words that do not necessarily convey meaning. But not all stop-words in English are removed from ISL. For this purpose, we created our own list of stop-words for English based on inputs from translations of sentences in our data set by expert ISL signers.

Among the words removed, the case of word sense disambiguation comes into play. Let us consider the word *do* which can appear in the following forms:

1. Do you like apples?
   In this sentence, *do* is a question word.

2. He does his homework. In this sentence, *do* is used as the main verb.

3. I do not like ice cream. In this sentence, *do* is an auxiliary to the verb *like*.

In the first and last sentence *do* needs to be removed from the final translation as it does not convey any extra meaning. However, in the second sentence, it is the main verb and hence conveys meaning and cannot be removed.

Following the removal of the stop-words, we convert all the remaining tokens into their root forms. At this stage, the tokens in the sentence represent a meta-language which we call as Text for ISL. It is essentially a representation of the signs that need to be performed in order.

4.3.3.4 Video Translation

The flowchart for converting Text for ISL into videos is described in 4.11.

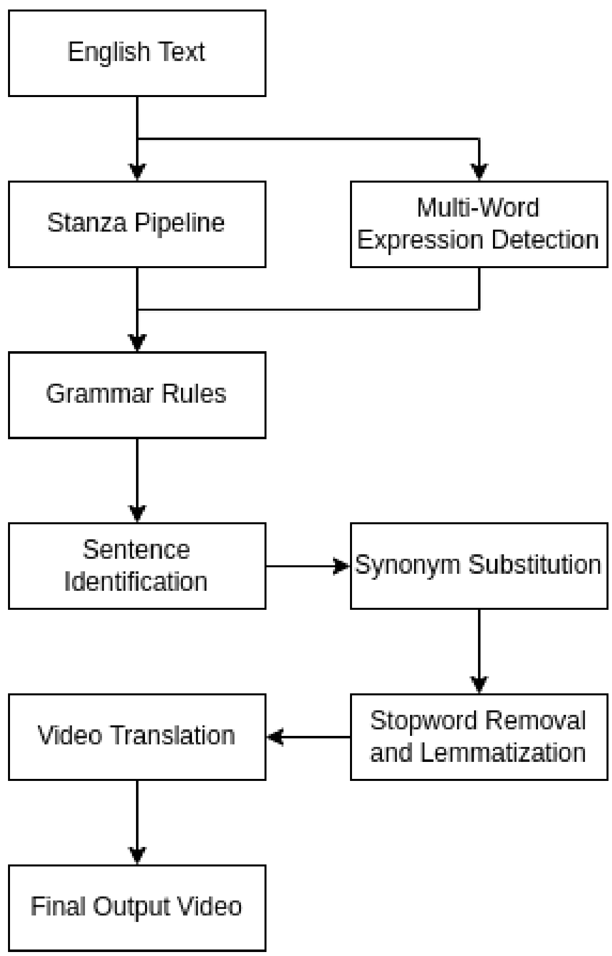


Figure 4.10: Englsih text to *text for ISL*

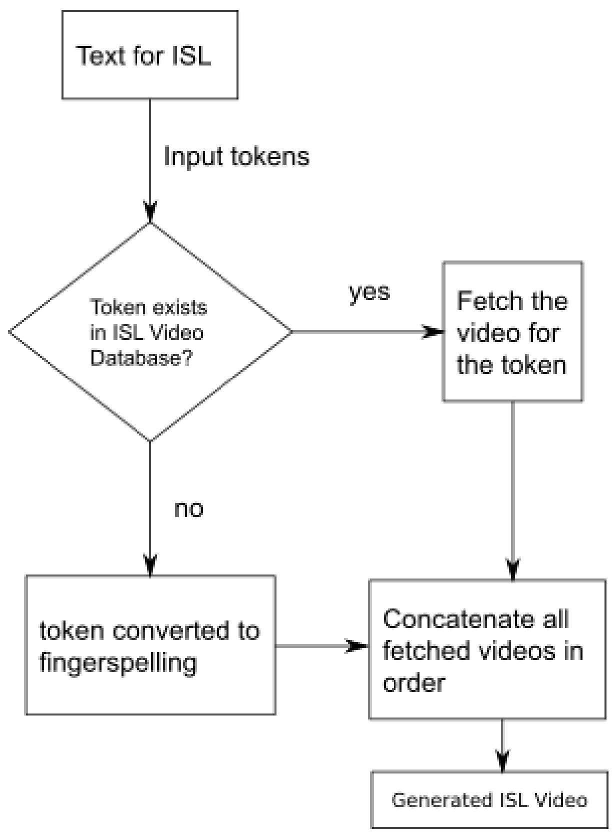


Figure 4.11: Converting *text for ISL* into videos

Chapter 5

Results and Discussion

Usually, for Machine Translation work, the accepted norm for evaluating results is the BLEU score [22]. The BLEU score aims to automate evaluating the correctness of translations by using n-gram counts in the target text and comparing it with the output text to judge the correctness of a sentence. This method however requires the availability of parallel bilingual data which is something which has not been done in the case of Indian Sign Language. We have to therefore go for manual evaluations. But manually evaluating 100,000 sentences is a nearly impossible task given the low number of experts in the language, most of whom are educators and have full-time jobs and cannot be realistically expected to evaluate the sentences manually.

So to evaluate the comparative accuracy of our model, we make some baseline assumptions.

1. We assume that the Stanford dependency parser is highly accurate since we are working with simple sentences.

2. We assume that the grammar rules applied by us are always true. The accuracy of the rules is not a matter of debate as they have been verified by both experts and other papers on Indian Sign Language. [11][31][8]

There have been works done in the rule-based translation of Indian Sign Language which we have discussed previously. So the only things left for us to evaluate are not the application of the rules but the other modules such as Multi-Word Expressions and Synonym Substitution which are the novel ideas introduced in this task.

5.1 Synonym Substitution

In our experiments, we found that out of 5592 sentences, 3821 sentences had words that were not in the ISL dictionary and could not be substituted by synonyms from the dictionary. We were able to substitute 1368 words with synonyms in 1011 sentences in total which is a significant improvement. Some of the substitutions were very accurate. Such as raise→lift, argument→debate, survive→last, astound→amaze, etc. However, not all substitutions were meaning conserving e.g. sword→steel, offend→break. Sometimes the substitutions were wrong because the meaning of the sign and the word were totally different. For example in some sentences, say became 'state' but the sign for 'state' is for the noun state(country) and not the verb state which actually would have shared the same meaning. The vocabulary lacks terms to words regularly used, such as newspaper, inspect, several, splash, gulp, goodwill, etc. Table 5.1 shows some examples where synonym substitution did take place. As we can see, not all substitutions were good substitutions such as sentence 1.

Sentence	After Grammar Rules	Words not in ISL dictionary	Replacement Word	Final Output
Niesiecki himself mentions it.	Niesiecki himself it mentions	mentions	name	N I E S I E C K I himself it name
A woman sells newspapers	Woman newspaper sells	newspaper	paper	woman paper sell
Helene let her fall	Helene her let fall	let	allow	H E L E N E mine allow fall
This foreshadows later events	This events later foreshadows	foreshadow	predict	this event late predict

Table 5.1: Examples of sentences where synonyms were substituted

5.2 Multi-Word Expressions

422 MWE were identified in 5592 sentences. All of the MWEs identified were more or less correct. Some of the correctly identified multi-word expressions include look_up_to, rolling_stock, you_know, thank_you, etc. but the correctness of multi-word expressions and where to use them and where not to use them is something that the system does not take into consideration. For example, all_around is a MWE in our database but when we want to translate *We have offices all around the globe*, all_around does not have an associated term in the ISL dictionary but *around* and *all* have. So it would have been better to not apply the grouping in this case.

5.3 Video Translation

To generate the sign language videos from the final output of our texts, we simply concatenate the videos from the ISLRTC dictionary together. There are other methods which work better (using SiGML and HamNoSys) but our goal was to improve the grammatical accuracy of the outputs so not having those systems does not affect the final accuracy of our system.

From the final videos, we realize that the system lacks in representing complex animation and facial expressions of ISL signs, as there are limited facial/body expressions in HamNoSys notation for representing non-manual signs. Moreover, in ISL, there are different criteria for handling directional signs based on their context. Hence, depending on the direction, there can be a change in the position of a sign from beginning to end than its usual sign. The proposed system does not handle the indexing of signs based on context.

Chapter 6

Conclusions and Future Work

We try to tackle more complicated sentences here, which are simple in structure, but some of the meanings conveyed by them are complicated, and the previously used methods fall short in some regards. Therefore, we have used approaches such as Multi-Word Expression detection, Synonym substitution, and Co-reference Resolution to handle hidden meanings within the sentence structure. We needed simple English sentences to test out our algorithm, as the rules worked best for simple sentences.

We can use machine learning approaches to do the translations on a large dataset to convert the standard English sentences. The methods mentioned in the paper could be a way to convert a large dataset of sentences on which Deep Learning models such as sequence2sequence and transformers can be trained.

Related Publications

- **Abhigyan Ghosh** and Radhika Mamidi. 2022. English To Indian Sign Language: Rule-Based Translation System Along With Multi-Word Expressions and Synonym Substitution. In Proceedings of the 19th International Conference on Natural Language Processing (ICON), pages 123–127, New Delhi, India. Association for Computational Linguistics.

Other Publications

- Salil Aggarwal, **Abhigyan Ghosh**, and Radhika Mamidi. 2020. SUKHAN: Corpus of Hindi Shayaris annotated with Sentiment Polarity Information. In Proceedings of the 17th International Conference on Natural Language Processing (ICON), pages 228–233, Indian Institute of Technology Patna, Patna, India. NLP Association of India (NLPAI).

Bibliography

[1] A. Abbi. 1992. india as a linguistic area revisited. *Language Sciences*, 12(2):107–316, 1991.

[2] E. Aboh, R. Pfau, U. Zeshan, et al. 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, 2005.

[3] A. Akbik, D. Blythe, and R. Vollgraf. Contextual string embeddings for sequence labeling. In *Proceedings of the 27th International Conference on Linguistics*, pages 1638–1649, Santa Fe, New Mexico, USA, Aug. 2018. Association for Computational Linguistics.

[4] B. Bergman and Östen Dahl. *Ideophones in Sign Language? The place of reduplication in the tense-aspect system of Swedish Sign Language*, pages 397–422. De Gruyter Mouton, 2011.

[5] A. Braffort, M. Filhol, M. Delorme, L. Bolot, A. Choisier, and C. Verrecchia. Kazoo: a sign language generation platform based on production rules. *Universal Access in the Information Society*, 15:541–550, 2016.

[6] C. Brewer. The second edition of the oxford english dictionary. *The Review of English Studies*, 44(175):313–342, 1993.

[7] W. Coster and D. Kauchak. Simple English Wikipedia: A new text simplification task. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 665–669, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.

[8] T. Dasgupta, S. Dandpat, and A. Basu. Prototype machine translation system from text-to-Indian Sign Language. In *Proceedings of the IJCNLP-08 Workshop on NLP for Less Privileged Languages*, 2008.

[9] T. Dozat and C. D. Manning. Deep biaffine attention for neural dependency parsing. *CoRR*, abs/1611.01734, 2016.

[10] S. D. Fischer. Two processes of reduplication in the american sign language. *Foundations of Language*, 9(4):469–480, 1973.

[11] L. Goyal and V. Goyal. Automatic translation of english text to indian sign language synthetic animations. In *Proceedings of the 13th International Conference on Natural Language Processing*, pages 144–153, 2016.

[12] T. Hanke. Hamnosys-representing sign language data in language resources and language processing contexts. In *LREC*, volume 4, pages 1–6, 2004.

[13] J. Hutchins. Machine translation: A concise history. *Computer aided translation: Theory and practice*, 13(29-70):11, 2007.

[14] J. Jepson. Urban and rural sign language in india. *Language in Society*, 20(1):37–57, 1991.

[15] P. Kar, M. Reddy, A. Mukherjee, and A. M. Raina. Ingit: Limited domain formulaic translation from hindi strings to indian sign language. *ICON*, 52:53–54, 2007.

[16] A. Kato, H. Shindo, and Y. Matsumoto. English multiword expression-aware dependency parsing including named entities. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 427–432, Vancouver, Canada, July 2017. Association for Computational Linguistics.

[17] N. Kulkarni and M. Finlayson. jMWE: A Java toolkit for detecting multi-word expressions. In *Proceedings of the Workshop on Multiword Expressions: from Parsing and Generation to the Real World*, pages 122–124, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.

[18] E. Loper and S. Bird. Nltk: The natural language toolkit, 2002.

[19] M. Marcus, B. Santorini, and M. A. Marcinkiewicz. Building a large annotated corpus of english: The penn treebank. 1993.

[20] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller. Introduction to wordnet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244, 1990.

[21] J. Miller, J. Heilmann, A. Nockerts, A. Iglesias, L. Fabiano-Smith, and D. Francis. Oral language and reading in bilingual children. *Learning Disabilities Research and Practice*, 21:30 – 43, 01 2006.

[22] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. Bleu: A method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, page 311–318, USA, 2002. Association for Computational Linguistics.

[23] R. Pfau and M. Steinbach. Pluralization in sign and in speech: A cross-modal typological study. 10(2):135–182, 2006.

[24] C. Proctor, D. August, M. Carlo, and C. Snow. The intriguing role of spanish language vocabulary knowledge in predicting english reading comprehension. *Journal of Educational Psychology*, 98:159–169, 02 2006.

[25] P. Qi, T. Dozat, Y. Zhang, and C. D. Manning. Universal dependency parsing from scratch. *CoRR*, abs/1901.10457, 2019.

[26] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. D. Manning. Stanza: A python natural language processing toolkit for many human languages, 2020.

[27] S. Roy, A. K. Maiti, I. Ghosh, I. Chatterjee, and K. Ghosh. A new assistive technology in android platform to aid vocabulary knowledge acquirement in indian sign language for better reading comprehension in l2 and mathematical ability. In *2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 408–413. IEEE, 2019.

[28] S. Sinha. *Indian Sign Language: An Analysis of Its Grammar*. Gallaudet University Press, 2017.

[29] S. Song, M. Su, C. Kang, H. Liu, Y. Zhang, C. McBride-Chang, T. Tardif, H. Li, W. Liang, Z. Zhang, and H. Shu. Tracing children's vocabulary development from preschool through the school-age years: An 8-year longitudinal study. *Developmental Science*, 18, 06 2014.

[30] W. C. Stokoe. Sign language structure: An outline of the visual communication systems of the american deaf (= studies in linguistics, occasional papers 8). *Buffalo, NY: University of Buffalo*, 1960.

[31] Sugandhi, P. Kumar, and S. Kaur. Sign language generation system based on indian sign language grammar. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 19(4), apr 2020.

[32] M. Vasishta, J. Woodward, and S. De Santis. *An introduction to Indian sign language:(Focus on Delhi)*. All India Federation of the Deaf, 1980.

[33] B. Vauquois. A survey of formal grammars and algorithms for recognition and transformation in mechanical translation. In *Ifip congress (2)*, volume 68, pages 1114–1122, 1968.

[34] T. Veale, A. Conway, and B. Collins. The challenges of cross-modal translation: English-to-sign-language translation in the zardoz system. *Machine Translation*, 13(1):81–106, 1998.

[35] V. K. Verma and S. Srivastava. A perspective analysis of phonological structure in indian sign language. In *Proceedings of First International Conference on Smart System, Innovations and Computing: SSIC 2017, Jaipur, India*, pages 175–180. Springer, 2018.

[36] J. Woodward. The relationship of sign language varieties in india, pakistan, and nepal. *Sign Language Studies*, (78):15–22, 1993.

[37] U. Zeshan. *Sign Language in Indo-Pakistan: A description of a signed language*. John Benjamins, 2000.

[38] U. Zeshan. Interrogative constructions in signed languages: Crosslinguistic perspectives. *Language*, 80(1):7–39, 2004.

[39] U. Zeshan. Sign languages. In M. S. Dryer and M. Haspelmath, editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013.

[40] U. Zeshan and M. M. Vasishta. Sethna m. 2004,"implementation of indian sign language in educational settings"-volume 15, number 2. *Asia Pacific Disability Rehabilitation Journal*, pages 15–35.

[41] I. Zwitserlood, M. Verlinden, J. Ros, S. Van Der Schoot, and T. Netherlands. Synthetic signing for the deaf: Esign. In *Proceedings of the conference and workshop on assistive technologies for vision and hearing impairment, CVHI*, 2004.