A Framework for the Automatic Generation of Indian Sign Language

Tirthankar Dasgupta,[1] Anupam Basu,[2] Plaban Kumar Bhowmick[2] and Pabitra Mitra[2]

[1]Society for Natural Language Technology Research, Kolkata, India;
[2]Indian Institute of Technology, Kharagpur, India

ABSTRACT: Sign languages are visual-spatial natural languages having their own set of vocabularies, syntax and grammatical structures. Yet, designing machine translation system between spoken and sign languages generates a number of interesting challenges, mainly due to the difference in structure as well as modality of the source and target language pairs. This paper presents a rule-based framework for the English-To-Indian Sign Language machine translation system. The system can be used to disseminate information to the deaf people in India. The present system takes an English sentence as input, performs syntactic analysis, and represents the source language into an intermediate case frame representation. Finally, transfer grammar rules are used to generate the corresponding ISL sentence structure. As ISL does not have any written form, the final sentence is synthesized in terms of pre-recorded video streams. The prototype system has been primarily evaluated by the native signers of India. Our initial evaluation result shows a significant improvement over the existing baseline system.

KEYWORDS: Indian Sign Language morphology, machine translation, recursive case-frame, sentence generation, system evaluation

1. INTRODUCTION

Indian Sign Language (henceforth called, ISL) is the native language commonly practiced by the deaf community of India. It is a non-verbal communication that allows

---

Correspondence: Tirthankar Dasgupta, Society for Natural Language Technology Research, Kolkata, India; e-mail: {iamtirthankar@gmail.com}; {anupambas@gmail.com;} {plaban@ gmail.com}; {pabitra@gmail.com}

deaf individuals to convey thoughts and ideas using their hands, arms, and facial expressions. Unlike spoken languages, ISL uses gestures instead of sounds to express a thought. Despite common misconceptions, recent linguistics research established that ISL, along with other sign languages (SL), is a complete natural language having its own grammatical structures, phonology, and complex set of morphological properties (Stokoe 1960; (Zeshan 2003). The morphology is complex in the sense that, it exhibits both sequential as well as simultaneous affixation of its manual as well as non-manual components (Liddell & Johnson 1989).

Estimates are that more than 1 million deaf adults and around 0.5 million deaf children in India use ISL as a communication mode (Zeshan 2003). One of every five deaf persons in the world uses ISL as a mode of communication.

Due to the cross-modal nature of a SL, both SL and spoken language users face a huge communication gap while exchanging information. The advents of modern computer science and technology have provided some significant solutions to the abovementioned problems. The most promising among them is the automatic SL machine translation (MT) systems. Over the past few decades since their introduction, automatic MT systems have proven to significantly reduce the communication and comprehension problems between SL and spoken language communities.

The advancement of Artificial Intelligence and Natural Language Processing technologies, along with the need for building tools related to Indian sign language automatic machine translation, was the primary motivation behind conducting the research described in this paper.

Here we present a framework for the syntactic transfer of English text to ISL. Our approach uses transfer grammar rules for the ISL sentence generation. The prototype machine translation system developed is a unidirectional system performing a structural transfer between English and ISL. We define structural transfer as the transformation of the source representation, reflecting the structure of the source language, to a target representation. The transfer involves mapping between lexical entries and transfer grammar rules. The rules identify the choice of the entries from the lexicon. The final target structure is achieved by the addition or deletion of words and restructuring of source representation.

We chose English as our input language because of the unavailability of Indian language tools like *parser, POS tagger*, and *morphological analyzer*. Further, most deaf institutions in India, provides ISL resources in English language only.

The target users of this system are those native deaf people who do not have any knowledge of spoken or written languages like English. Such individuals learn ISL from their birth through interaction with deaf family members or in deaf school. ISL is used not only by deaf people but also by hearing parents of deaf children, hearing children of deaf adults, and hearing deaf educators (Zeshan et al. 2004). Therefore, the system can also be used as an educational tool for those who wish to learn ISL as their second language.

2. INDIAN SIGN LANGUAGE LINGUISTIC ISSUES

Recent SL linguistic research shows that the underlying principles of structure and organization for SLs are similar to spoken languages (Liddell & Johnson 1989). Both have a lexicon that comprises gestures and conventional signs along with a set of rules governing the usage of the lexical items. A signer often uses the three-dimensional spaces (called the signing space) around their body to describe an event. The signing space is classified horizontally, vertically, and laterally into 27 cubical regions as shown in Fig. 1.

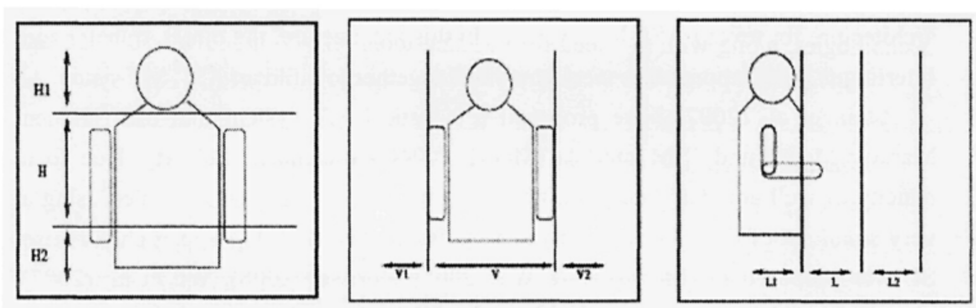

Fig. 1: Classification of signing space into horizontal, vertical, and lateral regions

The signs are made up of smaller meaningless formational units (like hand shape, location, orientation, movements and non-manual actions) that are linguistically significant, like the phonemes in words of spoken language (Sinha 2007). Each sign formational unit follows some definite constraints that are as complex as those of the spoken language. The ISL word order is relatively free and depends upon pragmatic factors (Zeshan 2003). The Subject-Object-Verb pattern, however, is preferred in most of the cases. For example, the English sentence "*He went to the park*" will be represented in ISL as:

`INDEX-IPSI PARK GO <PAST>`

Here, `INDEX-IPSI` is the ISL representation of the pronoun "He".

3. RELATED WORKS

In this section, we will discuss about some of the popular foreign SL MT systems. The ViSiCAST translator is an English-to-British Sign Language (BSL) translation tool (Safar & Marshall 2001; Marshall & Safar 2001). The system uses HPSG (Bengam et al. 2000) to represent source text into BSL. This system is one of the most successful developed so far. The TEAM project is a Text-To-American Sign Language (ASL) translation system, where the STAG formalism is used to represent source text into ASL syntactic structure (Zhao et al. 2000) The South African SL (SASL) MT system uses STAG grammar to represent the source text into an intermediate form. The system is under development stage and uses a small set of transfer rules for sentence generation. Huenerfauth (2006) has proposed a multi-path architecture for text to ASL MT systems. In this architecture, the direct, transfer and Interlingua based approaches were combined together to build an ASL MT system.

Stein et al. (2007) have proposed a statistical MT system that uses Hidden Markov Model and IBM models (Moore 2004) for training the data. Due to a paucity of well annotated corpora, however, the system has been evaluated using a very small set of data. Further, an Example Based MT (EBMT) approach for Irish SL was proposed in (Morrissey & Way 2005; Morissey 2008). Wu et al. (2007) proposed a Chinese SL MT system that uses a hybrid transfer based statistical model for translating Chinese text to Taiwanese SL (TSL).

None of the systems described above can be directly used for ISL MT because sign languages over the world are not universal. Every country has its own SL with varying grammatical structure and vocabulary. Secondly, most systems were short-lived with varying degree of success and hence cannot be considered for extension to ISL. Finally, most systems are domain dependent.

In India, ISL MT is still in its nascent stage. The only ISL MT system known so far is the INGIT system. INGIT is a Hindi-To-ISL MT system that has been built for the railway reservation domain (Kar et al. 2007). The system takes Hindi text input from the reservation clerk and translates into ISL. The output of the system is an animated representation of the ISL-gloss strings.

4. OVERVIEW OF OUR APPROACH

The architecture of our system is shown in Fig. 2. The main components of a translation system are (a) Analysis, (b) Intermediate Representation, (c) Transfer, and (d) Generation and e) Synthesis. We will discuss each component in the following sections.

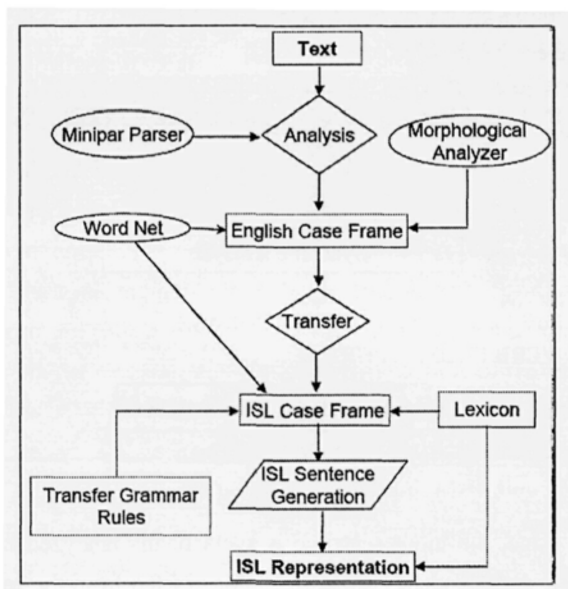

Fig. 2: System Architecture of the English to ISL Machine Translation System

5. INPUT SENTENCE ANALYSIS

The input text is first passed to the preprocessing unit, where we try to identify the frozen phrases. We prepare a hash table consisting of 350 frozen phrases and temporal expressions. The input sentence is passed to the phrase analysis module, where the occurrence of any frozen phrases or temporal adjuncts is identified and the positional index marked. If a sentence contains any frozen phrases, then all are replaced by some very uncommon noun word; the position of the occurrence of each frozen phrase is marked separately. The modified sentence is then passed to the parsing module for the syntactic analysis. Finally, before the intermediate representation of the source text, this noun word along with its positional index is replaced by the corresponding frozen phrase. The algorithm to handle the frozen phrases is presented below:

```
if sentence contains FROZEN PHRASE then
if FROZEN PHRASE ∈ NOUN PHRASE
    replace NOUN FROZEN PHRASE by NOUN + Position
else if FROZEN PHRASE ∈ VERB PHRASE
    replace VERB FROZEN PHRASE by VERB
else if sentence contains TEMPORAL EXPRESSION then
    replace TEMPORAL EXPRESSION by TEMPORAL + POSITION
end if
end if

                    [ Syntactic Analysis ]

replace NOUN + Position by NOUN FROZEN PHRASE
replace VERB by VERB FROZEN PHRASE

                    [ Intermediate Representation ]
```

In the next stage, the input sentence is syntactically analyzed using the Minipar parser (Lin 1998). A dependency structure is constructed from the resultant parse tree that helps identifying the functional parameters of the sentence like subject, object, modifiers, and adjuncts along with the morphological and Part of speech (POS) information. Finally, a lexical and morphological analysis of the input text is performed, where the input sentence is first tokenized to words and for every word, the associated POS is identified. The POS of a word can be obtained from the parsed structure itself. Each word along with its POS is sent to an English morphological analyzer for the identification of morphological information like the root, tense, aspect, person, number, and gender. We chose to use the PC-Kimmo English morphological analyzer for our task because of its wide recognition (Antworth 1990).

The functional parameters of a sentence along with the morphological information of each word are then represented into a language independent intermediate recursive case-frame structure as discussed in the next section.

6. THE INTERMEDIATE REPRESENTATIONS

The basic building block of our intermediate representation is a case-frame. A frame may have one or more sub frames. The value of a frame may be a word, a phrase, or another frame, thus making it recursive in nature. This recursive representation enables simple and compound, as well as complex sentences to be handled easily. The present paper focuses only on simple sentences.

The dependency structure as generated by the dependency parser encodes the grammatical relation of a sentence like subject, object, and verbs. Our case-frame structure represents the higher syntactic and functional information of a sentence (commonly known as feature structure) by a set of attribute-value pairs. Thus, a frame contains an attribute and a value corresponding to the attribute. The attribute corresponds to the name of a grammatical symbol (e.g. NUMBER, TENSE, and PERSON) or a syntactic function (e.g. SUBJECT and OBJECT), and the value is the corresponding feature possessed by the concerning constituent.

The verb frame structure contains the logical meaning of the verb along with the associated grammatical information like the tense and aspect. The argument frame contains information regarding the verb arguments and functional units like subject, direct object, indirect object and adjuncts. This information is stored in the form of an attribute value pair along with the POS of each value. These functional units are further classified into recursive frames. Figure 3 shows the frame representation of an English sentence "*The president of India will visit the capital of West Bengal next week*".

7. THE TRANSFER PHASE

The transfer mechanism described here is a sequential mapping of source language frame (s-frame) to target language frame (t-frame). This transfer is done by a mapping function that uses a bilingual transfer lexicon to map the source lexical item to the target lexical item. Each value corresponding to the attributes of the ISL case frame is represented by ISL glosses. The gloss convention used is the representation of input text root word in uppercase.

The mapping function identifies the individual English sub frames separately and based on the lexical elements and the grammatical information of the source language, returns the corresponding ISL structure in the form ISL gloss notation. In some cases the source language lexical elements does not have a direct replacement of the target language. For example, word like "*dinner*" is replaced by "NIGHT FOOD" in ISL.

Another important task performed by the lexical selection module is the handling of wh-phrases. Most of the ISL wh-signs fall under a single class "WHAT". All other wh-phrases like, WHERE, HOW, WHO, and WHY can either be represented directly by the sign "WHAT" or can be derived from it as shown in Table 1.

$$\begin{bmatrix} Type & \{Assertive\} \\ Verb & \begin{bmatrix} Root & \{Visit\} \\ Aux & \{Will\} \\ Tense & \{FUTURE\} \\ Aspect & \{NEUTRAL\} \end{bmatrix} \\ Argument & \begin{bmatrix} Subject & \begin{bmatrix} PRED & President \\ Pre-Mod & The \\ Post-Mod & of\ India \end{bmatrix} \\ Object & \begin{bmatrix} PRED & Capital \\ Pre-Mod & The \end{bmatrix} \end{bmatrix} \\ Adjunct & \begin{bmatrix} Preposition & \begin{bmatrix} P.Type & \{of\} \\ P.Arg & [West\ Bengal] \end{bmatrix} \\ Temporal\ Adjunct & (Next\ Week) \end{bmatrix} \end{bmatrix}$$

Fig. 3: Case-Frame Representation of the Sentence "The President of India will Visit the Capital of West Bengal Next Week"

TABLE 1

English wh-word with ISL derived form

English wh-phrases	ISL Derived Form
WHO	FACE + WHAT
WHERE	PLACE + WHAT
WHICH	$INDEX_{CONTRA}$ + $INDEX_{FRONT}$ + $INDEX_{IPSI}$ + WHAT
HOW MANY	COUNT + WHAT

7.1 The Bilingual Transfer Lexicon

The bilingual lexicon contains a list of English word along with the part of speech (POS) information, and the corresponding ISL video sign. The ISL lexical elements are represented in the bilingual lexicon in the form of glosses as shown below,

Play_V → PLAY
Play_V_PAST → PLAYED

Here, "_V" represents "Verb". The gloss entry in the bilingual lexicon does not contain any phonetic descriptions regarding a sign. Further, we define a uniqueness constraint upon each lexical unit within the ISL lexicon. The uniqueness constraint will allow no ambiguity between entries selected by the transfer lexicon and the actual lexical element present in the ISL lexicon thus, removing any lexical ambiguities.

7.2 Morphology Generation

The ISL verb group is classified according to its *tense* and *aspect*. In ISL, the verb tense is classified as PAST, PRESENT, and FUTURE. There are several kinds of aspects like, *habitual, perfect, frequentative, progressive, inceptive, and distributive.* For the present system, however, we have considered only three kinds of aspects: progressive, perfect, and neutral. Table 2 lists the various ISL tense and aspects along with the rules of incorporating morphological markers. During the morphological generation of each verb, the generator first identifies the root verb form (from the ISL case frame structure), and then the tense and aspect markers are sequentially added to the verb. We illustrate this process with an example. Consider the sentence "*She cooked*

TABLE 2

ISL Verb Tense and Aspect Along with the Rules for Attaching the Markers

Tense	Aspect	Markers
Past	Neutral	Root+<finish>
	Progressive	Root+<repeat>+<before>
	Perfect	Root+<finish>+<before>
Present	Neutral	Root
	Progressive	Root + <continue>
	Perfect	Root + <finish>
Future	Neutral	Root + <after>
	Progressive	Root + <after>
	Perfect	Root + <after>+ <finish>

the food". After the morphological analysis of the verb "*cooked*", we get the ISL verb case frame as:

$$\left[\text{Verb} \begin{bmatrix} \text{Root} & \{\text{COOK}\} \\ \text{Tense} & \{\text{PAST}\} \\ \text{Aspect} & \{\text{Neutral}\} \end{bmatrix} \right]$$

From the above case frame representation and based on the generation rule as mentioned in Table 2, the final ISL gloss representation of the verb "*cooked*" will be COOK+<FINISH>. Similar to the verb morphology, ISL noun morphological inflections are also represented by markers. Here, plural nouns are attached with the marker <*many*>. However, singular nouns do not have any morphological attachments. For example, the plural "Brothers" will be represented in ISL as BROTHER+<*many*>.

In ISL, pronouns are represented by Indexical signs (denoted as INDEX). Indexical signs are used to locate objects to a particular position in the signing space to establish a relationship between them (Zeshan 2003; Sinha 2007). The assignment of spatial location to the object is not entirely random, but rather depends upon the morphological features of the referred object like the *gender*, *person*, and *number*.

Table 3 shows the rules for representing different pronoun classes by indexical signs. From the table we can understand that, the ISL representation of the pronouns *He*, *She*, and *You* will be MASC-INDEX-IPSI, FEM-INDEX-IPSI and INDEX-FRONT respectively; MASC and FEM refer to the Masculine and Feminine markers.

TABLE 3

Rules for Representing ISL Pronouns by Indexical Signs

Number	Person	INDEX Attachments
Singular	1st	SELF
	2nd	FRONT
	3rd	IPSI/CONTRA
Plural	1st	SELF-ARC-SELF/ SELF-FRONT-SELF
	2nd	FRONT
	3rd	IPSI-FRONT-CONTRA

7.3 ISL Sentence Generation

To resolve the word order correspondence, the first stage of the transfer module is identifying the input sentence type. The sentence type is identified at the analysis phase. Based on the type of the input sentence, we define some generic transfer grammar rules for the final ISL sentence generation. Figure 4 lists 11 of the 36 transfer grammar rules that have been applied in our ISL sentence-generation

```
R1:Pred(X, Premod, Prep)->Pred (X)
R2:Pred(X, Premod, Indef Det)-> Pred
(X,Premod,INDEX (?))
R3:Pred(X, AUX)->0
R4:Pred(X, det)-> INDEX
R5:Pred(X, preplist)-> 0
R6:Tense(X,past),verb(X)->
     verb(X)+<FINISH>)
R7:[S_PM][S][S_M][V][V_M][O][O_M]A_T->
     A'_T[S'_M][S'_PM][O'_M][O'][V'_M][V']
R8:Aux[S][S_M][V][V_M][O][O_M]A_T->
A'_T[S'_M][S'][O'_M][O'][V'_M][V']Yes/No
R9:[Wh][S][S_M][V][V_M][O][O_M][A_T]->
[A'_T][S'_M][S'][O'_M][O'][V'_M][V'][Wh]
R10:Pred(X,pron,person=1,number=1)->
                      Pred (INDEX-IPSI)
R11: Pred (X, Premod, Adj)->
                   Pred (X,Postmod,Adj)
```

Fig. 4: Set of transfer grammar rules

module. We will illustrate the rules with an example. Consider the sentence, "*The director of NIHH will visit the deaf association of Kolkata in the month of January*".

We apply rule R4 to replace the determiners (Det) by the INDEX sign. Rule R5 is used to remove the prepositions from the sentence. The rule R6 specifies that, if the predicate X is a verb having a past tense, then attach the <FINISH> marker with the root form of the verb X. The word order of the output ISL sentence is handled by the rule R7. Here, S, S_{PM}, S_M are subject, it's pre-modifier and post-modifier. V, O, and V_M, O_M are the Verb, Object, and their modifiers. The attribute [X'] at the right hand side of the arrow in Rule 5, represents the ISL representation of the corresponding attribute [X]. On applying rules R4-7 we obtain the final ISL output structure as:

J-A-N-U-A-R-Y MONTH N-I-H-H DIRECTOR K-O-L-K-A-T-A DEAF ASSOCIATION VISIT+<FINISH>

Here, the proper nouns like, J-A-N-U-A-R-Y and N-I-H-H are represented by finger spelling. As ISL does not have any written form; the generated ISL is represented in the form of glosses.

8. SYNTHESES

As ISL does not have any written form, the final output of the system has to be a visual representation of the generated gloss. This visual representation may be done by several ways like (a) 2-dimentional pictures, (b) pre-recorded video signs, or (c) dynamically generated animations.

The 2-dimentional static pictures results in a loss of information particularly for signs having movement or non-manual components. This limitation can be solved by representing signs with pre-recorded videos. Both the figure and the video signs, however, suffer for morphologically inflected words. For example, the word "GO" has several inflected forms based upon its tense and aspect like GOING, GONE, and WENT. Image or video based representation must store all possible morphological inflections of a word in the dictionary, which not only needs a large storage space but also requires considerable time and effort to create such a large lexicon. This problem can be solved by using dynamically generated animations using Avatar-based technology. An Avatar takes the phonological information of a sign as input and generates the corresponding sign dynamically, thus eliminating the problems related to information loss, ambiguity, and storage space. Nevertheless, building such a large phonologically annotated lexicon is a non-trivial problem and requires much time and effort

Despite the disadvantages of huge storage requirement, our present sign synthesis module represents a sign by a concatenated stream of pre-recorded video files. We adopt this approach for implementing a demonstration system as well asfor evaluating the performance of our syntactic transfer rules.

9. SYSTEM EVALUATIONS

Evaluating a Text-to-ISL MT system is a difficult task. As ISL does not have any written orthography, standard techniques for evaluating Text-Text MT systems like BLEU (Papineni et al., 2002) and IBM Models (Moore, 2004) are not applicable for Text-to-ISL systems. Nonetheless, it is important that the ISL machine translation system should convey information that is understandable as well as acceptable to a deaf person. Hence, to identify a proper evaluation strategy for our system is essential.

The evaluation strategies of MT systems for natural languages are classified into two classes: (a) Automatic Evaluation and (b) Manual Evaluation. In automatic evaluation, the output of the system is compared with a set of gold standard data consisting of all possible correct translations of a given input sentence. The accuracy of the system is measured in terms of the degree of similarity between the output sentence and the gold standard data. As ISL does not have a standard written form, building such a large gold standard database for automatic evaluation requires a huge amount of time and effort. On the other hand, in user-based evaluation techniques, the output of a system is rated by several users. Based on the user's feedback, the performance of the system is calculated. Despite requiring much larger time and effort, user-based evaluation has a number of advantages over automatic evaluation as discussed in the literature (Huenerfauth 2006). Hence, we chose to perform a user-based evaluation for the present English to ISL MT system. In the next section we will discuss some of the issues that are to be handled before starting the evaluation process.

9.1 Issues Related to Evaluation

The users who are going to evaluate the system must be native ISL signers. A native ISL signer is a person who is deaf by birth and uses ISL as his first language from his childhood. This designation is important because there may be certain minute differences between the actual output and the translated output that only a native ISL user can identify. To understand the present system's features, however, we expect the ISL signers to have some knowledge of English. Hence, these ISL experts are not our target users. We created a corpus size of 247 sentences collected from 2 different sources. The first set contains 147 sentences collected from "*A-level introductory course in Indian Sign Language*", a work book published by the NIHH.[1] The second set contains 100 sentences collected from class four English textbook of West Bengal Primary School Education. We designed a GUI for evaluation purposes. The evaluator can watch the input sentence as well as the generated ISL representation as many times as he/she wants. Based on their observations, the evaluators assign a score to each output sentence.

9.2 Sign Understandability Test

The machine translated ISL output can either be represented in the form of images or pre-recorded videos. The image lexicon has a size of 1300 distinct entries, whereas the video lexicon contains only 350 entries. In the sign understandability test, we try to compare the performance of these two modes of representation (i.e image and video). We have randomly collected 100 ISL signs both in the form of image and video format. We then classify them into two classes: (a) one-hand signs and (b) two-hand signs. Each class contains a set of 50 signs. The signs are shown to a group of three ISL experts from NIHH, Kolkata who classify the signs as valid or invalid, according to their understandability and quality of the following metrics:

- Recognizing the hand shapes (HS)
- Recognizing finger & palm orientation (ORT)
- Recognizing hand location (LOC)
- Recognizing hand movements (MOV)
- Recognizing Non-manual articulations (NMA)

---

[1] Ali Yavar Jung National Institute for the Hearing Handicapped

Figures 5 and 6 summarize the comparative study between image and video based one handed and two handed signs. The X-axis specifies the different phonological parameters of a sign and the Y-axis shows the number of signs correctly recognized by an evaluator. From the study we can observe that representing signs with pre-recorded video performs better than image based signs. Most image

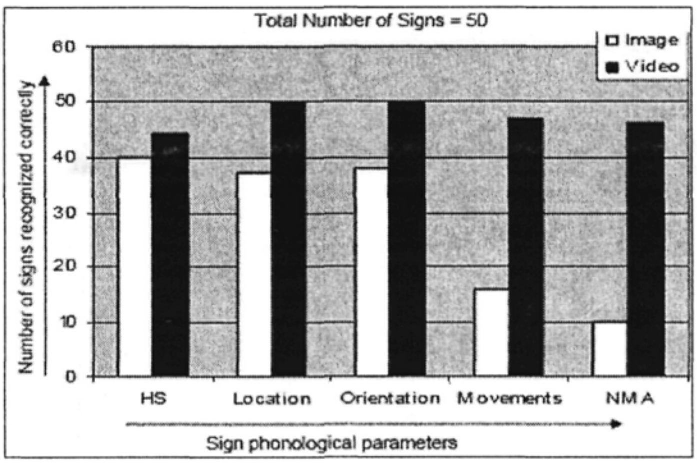

Fig. 5: Comparison between one hand image and video signs

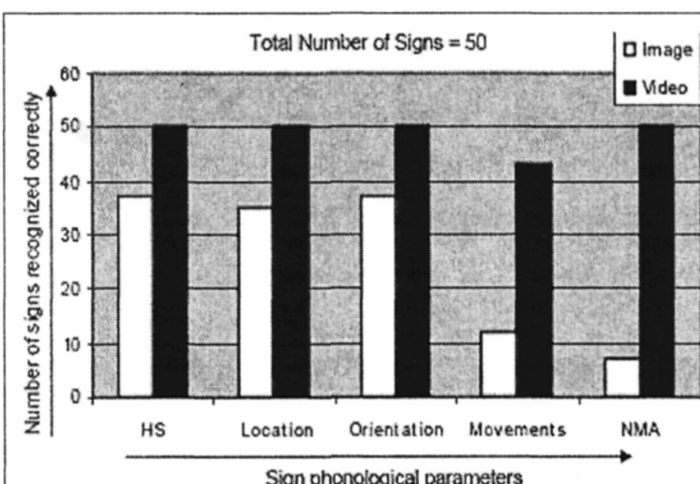

Fig. 6: Comparison between two hand image and video signs

signs are ambiguous as they fail to represent MOV and NMA. Figures 5 and 6, however, show that around 4% one-hand and 16% two-hand video signs failed to identify MOV component. This result is mainly due to the directional verbs where the movement varies depending on the location of subject, object, or both. Based on the above results, we choose to use the video lexicon in our present MT system despite having a much larger image database.

9.3 Measuring the System Performance

The machine translated ISL sentences were shown to the evaluators individually. Each evaluator rates the translated sentences based on two parameters *Intelligibility* and *Wellformedness*. The intelligibility and wellformedness is measured with the 4- and 6-point scale metrics, as proposed in the literature (Jordan et al. 2003). The Intelligibility of the system is computed using 147 test sentences from the test corpus of 247 sentences and the wellformedness test uses all the 247 test sentences from the corpus.

Table 4 defines the 4- and 6-point scale for intelligibility and wellformedness and shows number of sentences that have been assigned to each of the score value by the evaluators. Finally, we took the average score of each evaluator to get the overall accuracy of the system. The overall intelligibility and wellformedness score is shown in Figs. 7(a) and 7(b) respectively.

The Intelligibility score from Fig. 7(a) shows that around 74% sentences clearly understandable with some minor modifications like, removal of some alpha numerical characters, determiners, and prepositions that are not present in the stop word list, from the output. Among the Output sentences, 14% can be guessed due to minor word order and attachment problems, and around 8% sentences are totally incomprehensible. Out of this, around 5% sentences contain directional verbs and 3% sentences have compound constructions.

9.4 Comparing with the Baseline System

As the proposed prototype ISL MT system is a first of its kind in India, wecould not find any other system to compare our results. Hence, we choose the Sign Exact English (S.E.E) as our lower baseline. S.E.E represents each word of an English sentence by the corresponding ISL sign (Vasishta et al. 1973). The sentence is presented in original English word order without considering the ISL linguistic

TABLE 4

Intelligibility and Well-formedness metrics and number of sentences assigned to each class by the evaluators (E1, E2 and E3)

Score	Intelligibility (Total No. of Sentences = 147)	E1	E2	E3
1	Meaning is clear, no rewriting needed	76	65	74
2	Meaning is clear, need some rewriting	43	45	41
3	Meaning is not clear, can be guessed	20	22	21
4	Meaning is not conveyed at all	8	15	11
Score	Wellformedness (Total No. of Sentences = 247)			
1	Meaning is clear; grammar, word usage and style are all appropriate and no re writing is needed	12 3	12 2	12 3
2	Minor correction needed	32	29	28
3	Word order errors	38	34	32
4	Meaning is guessed, Attachment tense and number errors	12	21	25
5	Phrase and clauses missing	24	23	19
6	Subject and predicate missing	11	11	13

(a)

(b)

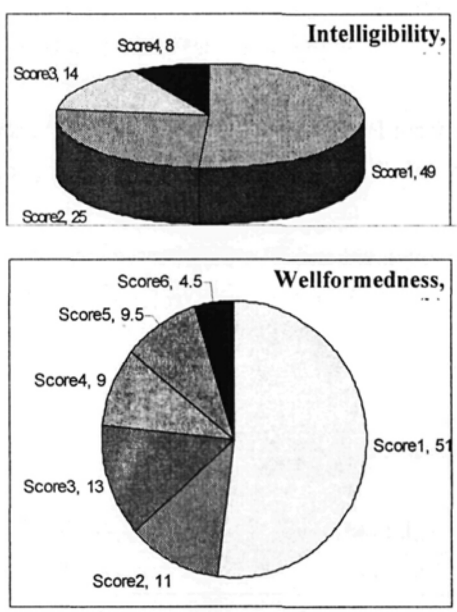

Fig. 7: Overall Intelligibility (a) and Wellformedness (b) Scores of the Present System (score values are shown in percentages)

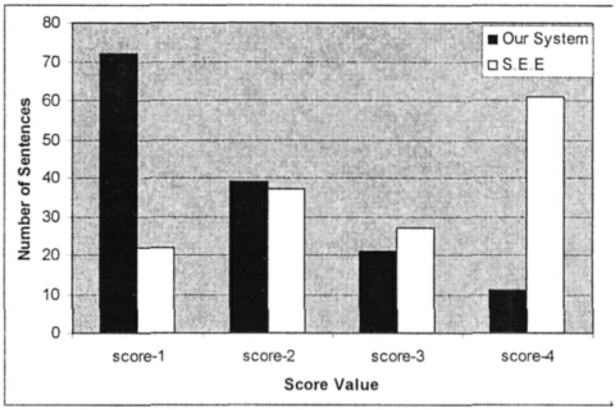

Fig. 8: Comparing Intelligibility between our system and S.E.E

constraints. The graph in Fig. 8 shows the comparison of the direct translated S.E.E system and our system.

9.5 Inter-Evaluator Agreement

In our present evaluation methodology, we used Cohen's Kappa (Cohen 1960) to measure the reliability in the evaluation data as provided by the three evaluators on intelligibility and well-formedness. The Kappa coefficient **K** is defined as,

$$\mathbf{K} = (P_o - P_e) / (1 - P_e)$$

where, P_e is the expected agreement and P_o is the observed agreement.

From the **K** value of both evaluation metrics (refer to Table 5), we conclude that the evaluation results provided by the evaluators are consistent and hence reliable.

TABLE 5

Inter Evaluator Agreement

	Intelligibility	**Well-formedness**
P_o	0.88	0.81
P_e	0.31	0.35
K	**0.83**	**0.71**

10. CONCLUSION AND FUTURE WORKS

Although, there have been a lot of improvements in MT theory and technologies, we are still a long way from automated ISL MT that will generate animated signers. In this paper we present a structure transfer framework for the English-To-ISL machine translation system that uses transfer grammar rules to generate the output ISL gloss structure. These glosses are finally represented by prerecorded video streams. The prototype system handles only English simple sentences. We have evaluated the system by native ISL signers and our initial results are promising. Based on the evaluator's feedback we have done some changes added more rules into the system. However, the work is by no means complete and the field still faces number of novel and non-trivial challenges. The most important among them is the need to display the output of such a system as animated sign. This will help to handle complex morphological features of a sign language like directionality, classifiers and index can easily be handled. Further, more rules are to be added to make the system generic and handle compound as well as complex sentences. In the next phase of our work we will try to address some of the above mentioned issues.

REFERENCES

Antworth, E.L. 1990. PC-KIMMO: a two-level processor for morphological analysis, *Technical report, TX Summer Institute of Linguistics.*

Bangham, J.A., Cox, S. J., Elliot, R., Glauert, J. R. W., Marshall, I., Rankov, S. and Wells, M.2000. Virtual signing: Capture, animation, storage and transmission—An overview of the ViSiCAST project, *IEEE Seminar on Speech and language processing for disabled and elderly people.*

Cohen, J. 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, **20**, 37-46.

Huenerfauth, M. 2006. *Generating American sign language classifier predicates for English-To-ASL machine translation.* PhD thesis, University of Pennsylvania.

Jordan, P.W., Dorr, B.J., Benoit, J.W. 1993. A first pass approach for evaluating machine translation systems, *Machine Translation*, **8**.

Kar, P., Reddy, M., Mukherjee, A. and Raina, A.M. 2007. INGIT: Limited domain formulaic translation from Hindi strings to Indian sign language, *International Conference on Natural Language Processing.*

Liddell, S. and Johnson, R.E. 1989. American sign language: The phonological base, *Sign Language Studies,* **64**, 195-277.

Lin, D. 1998. Dependency-based evaluation of MINIPAR, *Workshop on the Evaluation of Parsing Systems*, Granada, Spain.

Marshall, I. and Sáfár, É. 2001. Extraction of semantic representations from syntactic SMU link grammar linkages. *Proceedings of Recent Advances in Natural Language Processing*, 154-159, edited by Angelova G., Tzigov Chark, Bulgaria, September.

Moore, R.C. 2004. Improving IBM word-alignment model 1, *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 518, Barcelona, Spain

Morrissey S. and Way A. 2005. An example-based approach to translating sign language, *Proceedings of Workshop Example-Based Machine Translation (MT X-05)*, Phuket, Thailand.

Morrissey S., 2008. *Data-driven machine translation for sign languages*. PhD thesis, Dublin City University.

Papineni, T.W.K., Roukos, S. and Zhu, W. 2002. Bleu: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphia, 311-318.

Pollard, C.J. and Sag, I.A. 1994. *Head-driven phrase structure grammar*, Chicago, University of Chicago Press.

Sinha, S. 2007. *A skeletal grammar of Indian sign language*. PhD thesis.

Stein, D., Bungeroth, J. and Ney, H. 2006. Morpho-Syntax based statistical methods for sign language translation. *Proceedings of the 11th Annual conference of the European Association for Machine Translation*, Oslo, Norway.

Sáfár, É. and Marshall, I. 2001. The architecture of an English-text-to-Sign-Languages translation system. *Recent Advances in Natural Language Processing (RANLP)*, edited by Angelova G, 223-228

Stokoe, W.C. 1960. *Sign language structure: an outline of the visual communication systems of the American deaf.* Silver Spring, MD: Linstok Press.

Vasishta, M., Woodward, J. and DeSantis, S. 1985. *An introduction to the Bangalore variety of Indian sign language*. GRI Monograph, 4.

Vasishta, M., Woodward, J. and DeSantis, S. 1998. *An introduction to Indian sign language*, All India Federation of the Deaf (Third Edition).

Woodward, J. 1973. Some characteristics of Pidgin Sign English, *Sign Language Studies*, **3**, 39-46.

Wu, C.H., Su, H.Y., Chiu, Y.H. and Lin, C.H. 2007. Transfer-based statistical translation of Taiwanese sign language using PCFG. *ACM Transactions on Asian Language Information Processing* (TALIP), 9-16.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N. and Palmer, M. 2000. A machine translation system from English to American sign language. *Association for Machine Translation in the Americas*.

Zeshan, U. 2003. Indo-Pakistani sign language grammar: A typological outline, *Sign Language Studies*, **3**, 157-212

Zeshan, U., Vasishta, M. and Sethna, M. 2004. Implementation of Indian sign language in educational settings, *Asia Pacific Disability Rehabilitation Journal*, **15**, 15-35.