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

Tirthankar Dasgupta
IIT, Kharagpur
tirtha@
cse.iitkgp.ernet.in

Sandipan Dandpat
IIT, Kharagpur
sandipan@
cse.iitkgp.ernet.in

Anupam Basu
IIT, Kharagpur
anupambas@
gmail.com

Abstract

This paper presents a prototype Text-To-Indian Sign Language (ISL) translation system. The system will help dissemination of information to the deaf people in India. The current system takes English sentence as input, performs syntactic analysis, and generates the corresponding ISL structure. Since ISL does not have any written form, the output is represented in terms of pre-recorded video streams. The system uses Lexical Functional Grammar (LFG) formalism for representing ISL syntax.

1 Introduction

The All India Federation of the deaf estimates around 4 million deaf people and more than 10 million hard of hearing people in India (Zeshan et al, 2004). Studies revealed that, one out of every five deaf people in the world is from India. More than 1 million deaf adults and around 0.5 million deaf children in India uses Indian Sign Language (henceforth called ISL) as a mode of communication (Zeshan et al, 2004). ISL is not only used by the deaf people but also by the hearing parents of the deaf children, the hearing children of deaf adults and hearing deaf educators (Zeshan et al, 2004).

Due to their inability in accessing information through common broadcast modes like television, radio etc., and communication for the deaf community in common places like railway, bank, and hospitals is difficult.

Efforts to extend the existing means of communication for the hearing impaired include close circuit captioning in television and communication through interpreter. The first approach assumes a good knowledge in written languages like English, Hindi, or Bengali. The second approach is not always practically feasible.

A large section of the hearing impaired in India uses ISL as their mode of communication. However, due to the inherent difficulty in their written texts, an automatic Text-to-ISL translation system could help to make more information and services accessible to the hearing impaired. Moreover, the system will not only improve information access, but it can also be used as an educational tool to learn ISL.

Though some work has been done on machine translation (MT) from English to American or British Sign Language (SL) (Huenerfauth, 2003), but for ISL, MT systems are still in its infancy. The underlying architecture for most of the systems are based on:

I. Direct translation: This requires knowledge of both the source and the target language. Moreover, word order of the output may not be the desired one.
II. Statistical MT: It requires large parallel corpora which is very difficult to collect.
III. Transfer based architecture. As ISL does not relate to other SLs of either Asia or Europe (Zeshan, 2003), the existing systems transfer grammar rules cannot be applied to translate English to ISL.

Further, some of the systems are domain specific in nature, and cannot be used to generic systems. Hence, most of the above systems remain unusable for the deaf community of India. This is the prime motivation behind building a generic English Text-to-ISL translation system.

The objective of this paper is to present a prototype English-to-ISL generic machine translation

Proceedings of the IJCNLP-08 Workshop on NLP for Less Privileged Languages, pages 19–26,
Hyderabad, India, January 2008.

system. Currently the system takes simple English sentences as input and generates ISL-gloss which may then be converted into the Hamburg Notation System (HamNoSys)[1] (Prillwitz et. al, 1989). The HamNoSys representation will provide signing instructions to the sign synthesis module, to generate an animated representation of ISL to the user. Lexical Functional grammar (LFG) f-structure is used to represent ISL syntax.

The paper is organized as follows: Section 2 presents linguistic issues related to ISL. Section 3 presents a brief summery of the related works. Section 4 presents the overall system architecture. Section 5 presents system evaluation and results. Section 6 presents the sign synthesis via HamNoSys, and Section 7 presents conclusion and future work.

2 ISL Linguistic Issues

Indian Sign Language (ISL) is a visual-spatial language which provides linguistic information using hands, arms, face, and head/body postures. A sign is a sequential or parallel construction of its manual and non-manual components. A manual component can be defined by several parameters like hand shape, orientation, position, and movements where as non-manual components are defined by facial expressions, eye gaze, and head/body posture (Zeshan, 2003). However, there exist some signs which may contain only manual or non-manual components. For example the sign "Yes" is signed by vertical head nod and it has no manual component.

ISL lexicon is categorized according to the spatial behavior of the signs (Zeshan, 2003). There are three open lexical classes: i) Signs whose place of articulation are fixed, like, "hand", "teeth", "eye", "me", and "you" as shown in Fig. 1. ii) Signs whose place of articulation can change, like, "good," "friend," and "marry" as shown in Fig. 2. iii) Directional signs are those where there is a movement between two points in space. For example, in the sentence "I help him" the head word is "help" and direction of the sign is from subject "I" to the object "him" (Fig. 3). Directional signs generally show verbal property (Zeshan, 2003). Apart from the directional signs, ISL morphology is mostly derivational in nature and there are no affixes in signs. The closed lexical class contains classifier hand shapes, discourse markers, and non-manual signs (Zeshan, 2003). A classifier hand shape contains specification related to hand configuration that represents the characteristics of a referent. For example, consider the sentence "*Put the cup on the table*". Here the hand configuration will contain shape of a "*cup*" added with a movement to express the event "*put*".

ISL discourse structure is classified into manual and non-manual markers. Manual discourse markers can occur either in clause final position (as in, "*it's over, what else we can do*?") or in clause initial position (like, "*well, I have nothing to say*"). The non-manual marker like "head nodding" occurs only in clause final position after the last manual sign of the clause.

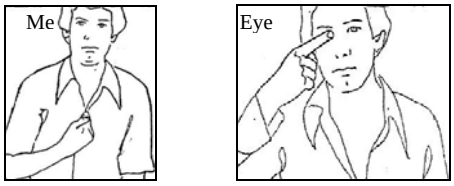

Fig.1: Signs whose place of articulation is fixed (Vasistha et. al 1998)

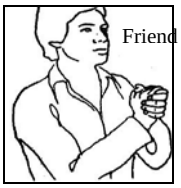

Fig. 2: Signs whose place of articulation can change (Vasistha et. al 1998)

Fig. 3: Directional Sign, "I help you". Taken from AYJNIHH workbook video CD.

3 The State-of-Art for Text-to-Sign Language

In spite of the advancements in modern computer science technology, there is a paucity of research in developing machine translation (MT) system on sign language particularly in India (Zeshan et al. 2004). Some of the MT systems for other sign lan-

[1] www.sign-lang.uni-hamburg.de/Projekte/HamNoSys

guage are briefly described below. The underlying MT architecture can be classified into i) Direct translation system, ii) Transfer based architecture and iii) Statistical MT.

The direct translation approach generates the SL by direct replacement of the words of input English sentence. Generally the word order of the SL remains the same as that of the English text. However, as in the case of English to ISL, the target SL may not allow the same word order. Also, the system assumes a strong knowledge of both the English as well as the target SL.

Some of the direct translation systems include:

- TESSA: A Speech-To-British Sign Language (BSL) translation system that aims to provide a communication aid between a deaf person and a Post Office clerk. The system uses formulaic grammar approach where a set of pre-defined phrases are stored and translation is done by using a phrase lookup table. However, the use of small set of sentences as templates makes TESSA a very domain specific system. It assumes a very restricted discourse between the participants (Cox, 2002).
- The SignSynth project (Grieve- smith 1998; Grieve-smith, 1999) uses ASCII-Stokoe model for the representation of Signs. The animated output is generated by converting ASCII-Stokoe into VRML (Virtual Reality Modeling Language). In his another project Grieve-Smith proposed a Text to American Sign Language (ASL) machine translation system. The system has been evaluated in the weather information domain.

In a transfer architecture system, the source language representation is transformed into a suitable syntactically/semantically correct target language form by applying proper transfer grammar rules. These rules are dependent upon both the source and the target language. However, as the source/target language changes new rules are need to be added. The transfer grammar approach is not only used in text to SL MT systems but also in text-to-text MT systems, like the Shakti MT system which is used to translate English text to Hindi (Bharati et. al., 2001; Bharati et. al., 2003). The transfer architecture systems include:

- The ViSiCAST translator, which is a English to British Sign Language (BSL) translation tool (Marshall & Sáfár, 2001; Bangham et al., 2000). The system uses HPSG (Pollard and Sag, 1994) formalism to represent source text into BSL and the grammar is implemented using a Prolog based system ALE. The system handles discourse phenomena by using Discourse Representation Structure (DRS) (Bos et. al, 1994) and the phonology is represented in HamNoSys. This is one of the most successful system developed so far (Huenerfauth, 2003).
- The ASL workbench (Speers, 2001) is a Text-To-ASL MT system which uses Lexical Functional Grammar (LFG) (Kaplan, 1989) formalism to represent English f-structure into ASL. The system uses a very sophisticated phonological model which is based on Movement-Hold principle of ASL phonology (Lidell & Johnson 1989).
- The TEAM project is a Text-To-ASL translation system where, the STAG (Synchronous Tree Adjoining Grammar) formalism is used to represent source text into ASL syntactic structure (Zhao et al, 2000). The system maintains a bilingual lexicon to identify the valid word-sign pair. The output of the linguistic module was a written ASL gloss notation. The manual and non-manual information, including the morphological variation, are embedded with in the ASL gloss notation. The output of the synthesis module uses animated human models (Avatar).

In addition, An Example based MT system for English-Dutch sign language was proposed by (Morrissey and Way, 2005). Stein et.al. (2006) has proposed a statistical MT system which uses Hidden Markov Model and IBM models for training the data. However, due to paucity of well annotated corpora, the system has been evaluated using a very small set of data.

3.1 Indian Scenario

INGIT is a Hindi-To-Indian Sign Language (ISL) Machine Translation system has been built for the railway reservation domain (Kar et. al, 2006). The system takes input from the reservation clerk and translates into ISL. The output of the system is an animated representation of the ISL-gloss strings via HamNoSys. INGIT is based on Hybrid-formulaic grammar approach unlike TESSA which uses purely formulaic approach. Here, Fluid Construction Grammar (FCG) (Steels and Beule, 2006)

is used to implement the Formulaic grammar. This is the only Hindi text-to-ISL machine translation tool encountered by us so far. However, the system is domain specific in nature and cannot be used for generic purpose. Further, the system does not have to handle any structural divergence between the source and the target language, as in most of the cases both Hindi and ISL show the same word order.

4 ISL MT Architecture

In order to overcome the above mentioned problem, we initially developed a direct translation system, however due to its inherent drawbacks, as mentioned in section 3, we need some other approach. One of the most popular techniques is to use statistical or case based MT system. However ISL does not have any written form, so it is very difficult to find any natural source of parallel corpora. Niedle et al. (2000) have proposed an approach to collect corpus for statistical MT research, in his approach first, annotation standard for the various hand shape movements was developed, then the Sign Language performances were recorded, and finally the recorded videos were manually transcribed. This is a very slow and expensive process. Due to the difficulty in obtaining parallel corpora of ISL, the statistical MT approaches may not be a feasible solution to our problem. Hence we decided to build a rule based transfer grammar MT system discussed in this section.

The system architecture of the proposed English Text-To-ISL MT system is composed of the following four essential modules (see Fig. 4):

1. Input text preprocessor and parser
2. LFG f-structure representation
3. Transfer Grammar Rules
4. ISL Sentence Generation
5. ISL synthesis

4.1 Text Analysis & Syntactic Parsing

The current Text-To-ISL translator takes simple English sentence as an input to the parser. We define simple sentence as, a sentence containing only one main verb. The input sentence is then parsed using the Minipar parser (Lin, 1998) and a dependency structure is constructed from the parse tree. However, before parsing, the input text is passed to the preprocessing unit, where we try to identify the frozen phrases[2] and temporal expressions[3] which the syntactic parser is unable to identify. We prepare a phrase lookup table consisting of 350 frozen phrases and temporal expressions which are identified before the input text is parsed. The parsing stage also includes classification of plural nouns. The plurality is identified using an English morphological analyzer.

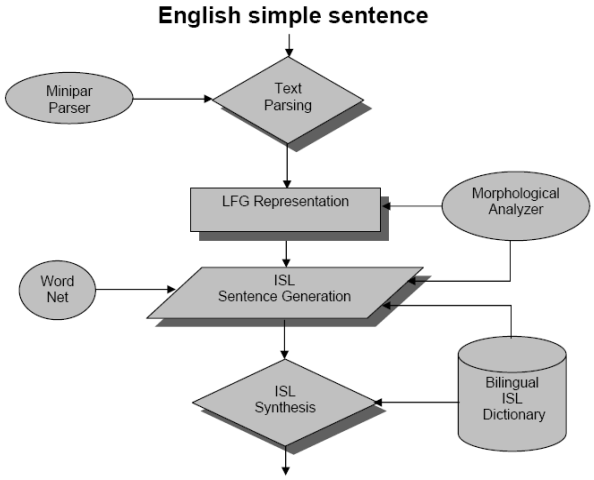

Fig. 4: Architecture of the Text-to-ISL MT system

4.2 LFG Representation

The Minipar generated dependency structure is more akin towards the LFG functional structure (f-structure). The f-structure encodes grammatical relation (like subject, object, and tense) of the input sentence. It represents the internal structure of a sentence. This includes the representation of the higher syntactic and functional information of a sentence. This higher syntactic and functional information of a sentence is represented as a set of attribute-value pairs. In an attribute-value pair, the attribute corresponds to the name of a grammatical symbol (e.g. NUM, TENSE) or a syntactic function (e.g. SUBJ, OBJ) and the value is the corresponding feature possessed by the concerning constituent. For example, Fig. 5 shows the attribute-value pair for the sentence "*John Played Cricket*". The main advantage of f-structure is in its abstract representation of syntactic and grammatical information of a sentence.

[2] Phrases that are composed of Idioms, and Metaphor
[3] Temporal Expressions contains Time, Day and Date.

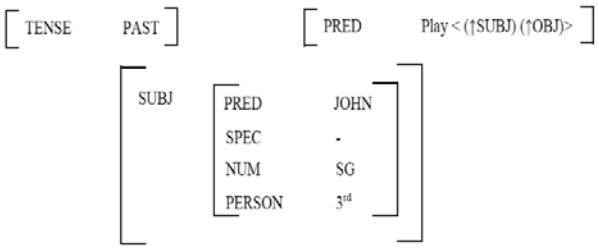

Fig. 5: Attribute-Value pair for the sentence "*John Played Cricket*"

4.3 ISL Generation

In the generation stage, English f-structure is converted to ISL f-structure by applying proper transfer grammar rules. Two main operations are performed during the generation phase: a) Lexical selection and b) Word order correspondence.
Lexical selection is done using an English-ISL bilingual lexicon. For example word like "Dinner" in English is replaced by "NIGHT FOOD" in ISL and "Mumbai" is replaced by the sign of "BOMBAY".

(1) English: "I had dinner with Sita"
ISL: "I SITA WITH NIGHT FOOD FINISH"

ISL has essentially a Subject-Object-Verb word order (unlike English which is Subject-Verb-Object). For Example, **(2)** shows the change in word order from English to ISL.

(2) **English**: "*I have a computer"*
ISL: "I COMPUTER HAVE".

However, in some cases the sign sentence depends upon the directionality of the verb as in **(3)**.

(3) **English**: "*I help you*"
ISL: "HELP + < hand movement from I-to-YOU>".

For sentences having only a subject and a verb, the subject always precedes the verb. Like:

(4) **English**: "*The woman is deaf*"
ISL*:* "WOMAN DEAF".

However, if the sentence contains a dummy subject **(5)**, then the subject is removed from the output.

(5) **English**: "*It is raining outside*"
ISL: "OUTSIDE RAINING"

For negative sentences, a negation mark is used after the verb **(6)**. The second bracket indicates a parallel non-manual component is attached with the sign "LATE".

(6) **English**: "*I am not late*"
ISL: "I {LATE + NOT}".

ISL has separate rules to handle adjectives occurring before a noun. In most of the cases an adjective must occur after the noun. However, if the adjective specifies a color then it should precede the noun (see **(7) & (8))**.

(7) **English**: "*The beautiful girl is playing*"
ISL: "GIRL BEAUTIFUL PLAY"

(8) **English**: "*I see a black cat*"
ISL: "I BLACK CAT SEE".

WH-Interrogative markers (like who, what, when, and why) always occur at the end of the sentence.

(9) **English:** "W*hen is your birthday?*"
ISL: "YOUR BIRTHDAY TIME+ QUESTION".

In case of yes/no type of questions, the sentence is followed by a non-manual yes-no marker (Zeshan, 2004).

(10) **English:** "Is the man deaf?"
ISL: "MAN {DEAF} yes-no"

Since ISL does not have any articles or conjunctions, they are removed from the generated output as shown in example **(2)-(10)**.

5 System Evaluation

Evaluating a Text-to-ISL MT system is difficult due to the absence of ISL written orthography. Hence, standard techniques for evaluating Text-Text MT systems are not applicable for Text-to-ISL systems. However, we have evaluated the system based on the feedbacks of the ISL experts. The generated outputs of the system are shown to the ISL experts and are classified as either valid or invalid according to their understandability and quality. The system was evaluated on a set of 208

sentences[4]. Table 1.1 summarizes the performance of the system. The overall system performance is around 90%. Most of the errors are due to compound sentences and directional verbs[5]. To understand the relative performance of the system on the simple sentences, we conducted two experiments removing compound construction and directional verbs. From the current experimental set up, 7% errors are propagated due to directional verbs and around 4% errors are due to compound constructions.

	No. of Sentences	Accuracy (%)
Overall Corpus size	208	89.4
Sentences without directional verbs	193	96.37
Sentences without compound constructions	201	92.53

Table1.1: Evaluation Results

6 ISL Synthesis

The ISL sentences thus generated are displayed via a stream of pre recorded videos or icons. However, it has been observed that the current approach of ISL synthesis is highly criticized (Grieve-Smith, 1999). As, representing ISL signs by pre-recorded video will result in loss of information related to discourse, classifier predicate, and directionality of sign. Also, storing sign video takes a lot of memory overhead. To overcome this crisis further developments are in progress. We represent ISL signs by HamNoSys and the generated HamNoSys string will be passed to the signing avatar.

6.1 HamNoSys

Sign language does not have any written form. In order to define a sign we need a notation system. The Hamburg sign language Notation system (HamNoSys) is a phonetic transcription system used to transcribe signing gestures. It is a syntactic representation of a sign to facilitate computer processing. HamNoSys is composed of several parameters by which a signing gesture can be defined like:

- Dominant hand's shape.
- Hand location with respect to the body.
- Extended finger orientation.
- Palm orientation
- Movements (straight, circular or curved)
- Non-manual signs.

Fig. 9 shows an example where HamNoSys representation of the word "WOMAN" is explained.

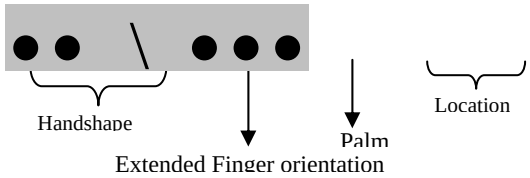

Fig. 9: HamNoSys representation of "WOMAN"

In this example, the parameters like movement and non-manual signs are not present, as the sign "WOMAN" in ISL does not have these expressions. Fig. 10 shows the ISL representation of "WOMAN".

Fig. 10: Sign of "WOMAN"
(Vashista et.al, 1998)

7 Conclusion and Future works

The paper presents a prototype text to ISL translation system. Our approach uses LFG f-structure to represent ISL syntax. As ISL does not have any written form, there is no standard source of ISL corpus. Hence, statistical MT methods are not feasible under such a condition. Our system is still under development stage. The sign synthesis module using an animated avatar has not been developed yet. We generate ISL output using pre-recorded ISL videos. Further morphological functionalities like, discourse, directionality, and classifier predicates are handled minimally

[4] Corpus collected from "'A' level Introductory course in Indian Sign Language" Work Book AYJNIHH.
[5] Verbs corresponding to directional signs.

In the next stage of our work, we will try to handle directional sign, discourse, and classifiers. The sign representation should be done using an animated avatar via HamNoSys notation. We will also develop the sign annotation tool and finally, a larger corpus will be built for a better evaluation and results.

References

N. Badler, R. Bindiganavale, J. Allbeck, W. Schuler, L. Zhao, S. Lee, H. Shin, and M. Palmer 2000. Parameterized Action Representation and Natural Language Instructions for Dynamic Behavior Modification of Embodied Agents. *AAAI Spring Symposium*.

J. A. Bangham, S. J. Cox, R. Elliot, J. R. W. Glauert, I. Marshall, S. Rankov, and M. Wells. 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.*

A. Bharati, D. M. Sharma, R. Sangal. 2001. AnnCorra : An Introduction, *Technical Report no: TR-LTRC-014*, LTRC, IIIT Hyderabad, Mar 2001, http://www.iiit.net/ltrc/ Publications/Techreports/TR-LTRC-14

A. Bharati, R. Moona, P. Reddy, B. Sankar, D.M. Sharma, R. Sangal, Machine Translation: The Shakti Approach, *Pre-Conference Tutorial at ICON-2003*.

J. Bos, E. Mastenbroek, S. McGlashan, S. Millies, M. Pinkal. 1994. A Compositional DRS-based Formalism for NLP Applications. *Report 59*. Universitaet des Saarlandes.

S. Cox, M. Lincoln, J. Tryggvason, M. Nakisa, M . Wells, M. Tutt, S. Abbott. 2002. Tessa, a system to aid communication with deaf people. *Fifth international ACM conference on Assistive technologies*.

M. Huenerfauth. 2003. A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems. *Technical Report MS-CIS-03-32*, Computer and Information Science, University of Pennsylvania.

A. Joshi, L. Levy and M. Takahashi. 1975. Tree Adjunct Grammar. *Journal of computer and system sciences*.

P. Kar, M. Reddy, A. Mukherjee, A. M. Raina. 2007. INGIT: Limited Domain Formulaic Translation from Hindi Strings to Indian Sign Language. *ICON*.

Ronald M. Kaplan. 1989. The formal architecture of lexical-functional grammar. *Journal of Information-Science and Engineering* 5: 305-322.

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

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

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

S. Morrissey and A. Way. 2005. An Example-Based Approach to Translating Sign Language. *In Proceedings of Workshop Example-Based Machine Translation* (MT X -05), Phuket, Thailand.

C. Neidle, J. Kegl, D. MacLaughlin, B. Bahan, and R. G. Lee. 2000. *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: The MIT Press.

C. J. Pollard, and I. A. Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press, Chicago, IL.

S. Prillwitz, R. Leven, H. Zienert, T. Hamke, and J. Henning. 1989. *HamNoSys Version 2.0: Hamburg Notation System for Sign Languages: An Introductory Guide, volume 5 of International Studies on Sign Language and Communication of the Deaf.* Signum Press, Hamburg, Germany,

É. Sáfár and I. Marshall. 2001. .The architecture of an English-text-to-Sign-Languages translation system.. *In G. Angelova, editor, Recent Advances in Natural Language Processing (RANLP)*, pp: 223-228. Tzigov Chark, Bulgaria.

G. Angus Smith. 1998. Sign synthesis and sign phonology. *Proceedings of the First High Desert Student Conference in Linguistics*.

G. Angus Smith. 1999. English to American Sign Language machine translation of weather reports. *Proceedings of the Second High Desert Student Conference in Linguistics.*

A. Speers. 1995. SL-Corpus: A computer tool for sign language corpora. Georgetown University.

A. Speers. 2001. *Representation of American Sign Language for Machine Translation*. PhD Dissertation, Department of Linguistics, Georgetown University.

L. Steels and J. Beule. 2006, Unify and Merge in Fluid Construction Grammar, *In: Lyon C., Nehaniv, L. & A. Cangelosi, Emergence and Evolution of Linguistic Communication, Lecture Notes in Computer Science*. Springe-Verlag: Berlin,.

D. Stein, J. Bungeroth and H. Ney. 2006. Morpho-Syntax Based Statistical Methods for Sign Language Translation. *In Proceedings of the 11th Annual conference of the European Association for Machine Translation*. Oslo, Norway.

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

Elizabeth Winston. 1993. *Spatial mapping in comparative discourse frames in an American Sign Language lecture*. Doctor of Philosophy in Linguistics diss., Georgetown University.

L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, and M. Palmer. 2000. A Machine Translation System from English to American Sign Language. *Association for Machine Translation in the Americas*.

U. Zeshan. 2003. Indo-Pakistani Sign Language Grammar: A Typological Outline. *Sign Language Studies - Volume 3*, Number 2 , pp. 157-212.

U. Zeshan. 2004. Interrogative Constructions in Signed Languages. *Crosslinguistic Perspectives Language - Volume 80*, Number 1, pp. 7-39.

U. Zeshan, M. Vasishta, M. Sethna. 2004. Implementation of Indian sign language in educational settings-Volume 15, Number 2, *Asia Pacific Disability Rehabilitation Journal*, pp. 15-35