Generating Indian Sign Language Text Using English/Hindi Text

Pawan Kumar, Savita Khatri

Assistant Professor, Dept of CSE, BMIET Sonipat, India

Abstract— **With the increasing need of the translation systems, research is being done in the Linguistics field and effort is being done in the development of the translators to assist hearing impaired people. Translation systems are able to convert an input into Indian Sign Language Text and then further into Sign Language symbols. Since, Indian Sign Language has its own syntax and grammar so a system is needed that can convert each sentence from English/Hindi grammar to ISL grammar.**

Keywords-component; Machine Translation, NLP, indian sign language, deaf, mute, ISL.

I. INTRODUCTION

DEAF and MUTE are the terms used for the people who are not able to speak or hear or in some cases both. These people are quite introverted due to the limited communication abilities and this poses quite a problem for the current world as the number of such people is increasing day by day. The only key solution to this communication problem is sign language.

Sign language is a form of non-verbal communication in which hands are used to make gestures to convey the message. Since sign language is known to only a limited number of people, this again creates a problem in communicating with the normal people. The non-trained people are not able to communicate with the deaf-mute people.

According to the 2011 census, there are 63 million people which sum up to 6.3% of the total population, who are suffering from significant hearing loss [7]. Out of these people, 76-89% of the Indian Deaf have no knowledge of language, either signed or spoken/written [8]. The reason behind the low literacy rate can be either of the following:

1) *Lack of Sign Language interpreters.*
2) *Unavailaility of ISL tool.*
3) *Lack of researches on ISL.*

Rest of the paper is organized as follows: section II describes the Indian Sign Language Grammar, section III specifies the review of previous research that has been done in this field, section IV discusses the challenges faced and finally, conclusion of paper is described in section V.

II. ISL GRAMMAR

Like other languages, Indian Sign Language has its own grammar. It is not dependent on the spoken language - English or Hindi. The sign language is not same as the manual representation of spoken English or spoken Hindi. It has certain unique and distinct features [2] like:

1. All the sign representation for numbers are done with appropriate hand gesture for every number. Eg. the sign for 45 will be the representation of four followed by sign representation of 5.
2. The signs for family relationships are preceded by signs for 'male/man' and 'female/woman'.
3. The interrogative sentences having words like WHAT, WHERE etc. are represented by placing these questions at the end of sentences.
4. The ISL consists of various non-manual gestures including mouth pattern, mouth gestures, facial expressions, body postures, head position and eye gaze.
5. The past, present and future tense is represented by the signs for before, then and after.

Due to their inability, communication for the deaf community in common places like railway, bank, and hospitals is difficult. To help them communicate better with the rest of the world, a system is needed which will enable the conversion of text to Indian Sign Language and vice versa. These systems will increase the quality of living of this community. Sign languages have not been studied as extensively as spoken languages, and there is still much left to be learned about them.

III. REVIEW OF PREVIOUS RESEARCH

There has been some amount of work done to promote communication between deaf- mute and normal people. Maximum work done is based on American Sign Language (ASL) or British Sign Language (BSL) but for Indian Sign Language (ISL), only a few systems are there for this purpose. These systems use Machine Translation as their underlying technique. Most of the systems using Machine Translation have the same underlying structure and are based on following three categories:

- *Direct Translation System*

In this system there is only word to word conversion. None of the context of sentences are taken into consideration. Words are directly transformed into target sign language without passing through an additional representation and without the change of grammar. In this translation, text is given as an input and target sign language is achieved without performing any kind of syntactic analysis on the original text. The word order i.e. the sequence of occurrence of words in the target sign language remains same as the word order of English text. But in the case of English to ISL, the word order of ISL may or may not be same as the given English/Hindi input text. So to overcome this problem a system is needed which has a strong knowledge of both English language as well as the target sign language. There is a "Domain Bounded English to Indian Sign Language

Translation Model" [6] that has been developed and it uses Direct Translation to convert English text to ISL text.

- *Transfer Based Translation*

In Transfer Based Translation systems, plain text is given as an input to the system. The system then analyses it syntactically and also semantically. It is then transferred into a sign language. In this system, source language is transformed into some intermediate abstract text, some linguistic rules are then applied to that text to transfer it into target language. Transfer based Translation is thus also known as "Rule based Translation" since some special set of rules are used which read the information of the source language and produce a semantic or syntactic structure in target language. This approach is used in text to sign language machine translation. A system named "TEAM" is used for text to ASL (American Sign Language) translation system [9].

- *Interlingua Based Translation*

In Interlingual system, a language independent semantic structure is produced by doing only semantic analysis on the source text. This independent structure is called Interlingua. The target language is then generated from this Interlingua. This system is an alternative of both direct translation system and transfer based system. One system named "ZARDOZ" is based on this translation which is used for English to Sign Language translation [10].

Only some work has been done specific to developing systems generating Indian Sign Language Text. Some of the systems developed using Machine Translation developed for this purpose are explained below:

A. INGIT

INGIT is a Sanskrit word which means signed. This system deals with the conversion of Hindi strings into Indian Sign Language. This system has been developed specific to Railway Enquiry Domain. It works on the strings of transcribed Hindi spoken language. A domain-specific construction grammar for Hindi is implemented in FCG. It first converts the input into a thin semantic structure. It is then given as an input

to ellipsis resolution which removes the unnecessary words. A saturated semantic structure is obtained after this step. Depending on the type of sentence (statement, query, negation, etc) ISL generator generates a suitable ISL-tag structure. This ISL tag structure is then passed to a HamNoSys converter to generate the graphical simulation. HamNoSys is a stoke based notation which are symbols used to represent sign language in written form.

The system was then validated by collecting data by interacting with the reservation counters. The data consisted of a vocabulary of 90 words which had all kind of words including verbs, related to time, specific to domain, numerals, months, cities, digits etc. They have shown some examples of ISL mappings which they have done. The mappings have either been handled at constituent level or they are formulaic mappings.

They have been successful in generating the semantic structure in about 60% cases.

B. Domain bounded English to Indian Sign Language Translation Model

This system [6] takes English text as input and converts it into ISL text which further gets translated into ISL symbols. The authors have developed the system for Railway Enquiries in India. So, the domain of this model is very specific. The architecture of the system is shown in figure 1. It has an input module which takes the input text to be translated. The tokenizer then splits the English sentence into separate words. They have made a repository of the ISL symbols only specific to railway enquiries. If a word does not have a sign associated in the repository, then the sign for the synonym of that word is used. A translator is used for this purpose of matching the word with the symbol. It filters the words to be translated by ignoring the words (for which there isn't any sign stored) and then accumulates the words in the sequence they were entered.

The drawback of this system is that it does not check for tenses and syntax and it is specific to Railway Enquiry only.
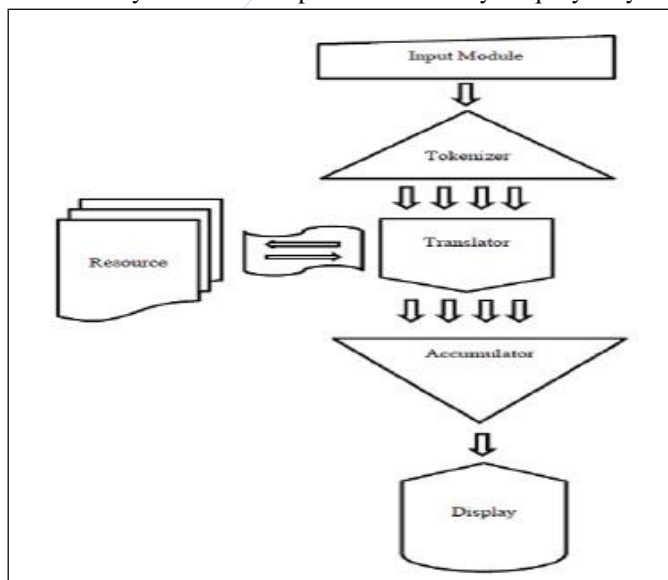

Figure 1. Architecture of the System [6]

C. Translator of Hindi Text to ISL and extension of ISL dictionary with WordNet

This system of Sign Language Generation [2] is divided into two phases. Phase A deals with the pre processing of Hindi Sentences and are converted into ISL grammar with the help of Dependency Parser and WordNet in combination. The Dependency Parser uses dependency graphs to represent words and their relationships between "head" words and words which modify those heads. For each word, the information is represented by the expression:

<word id> <word> <lemma> <POS Tag> <parent id> <dependency label>

WordNet is a system developed by IIT Bombay [16] which collaborate various types of relations between the Hindi Words from semantic to lexical forms.

Phase B deals with conversion of this grammar into Sign Language symbols using HamNoSys. Then SIGML is used to convert the symbols generated by HamNoSys into XML tags

form. This is then in a format readable by 3D rendering software

D. Prototype Machine Translation System

The system [5] generated take English language text as input and converts it into corresponding text with ISL structure. Their system shown in Figure 2 comprises of four main modules. They are: a) Text Analysis and Syntactic Parsing b) LFG f-structure representation c) Transfer grammar rules and d) ISL Sentence Generation. They have prepared a lookup table consisting of 350 frozen phrases and temporal expressions which are identified before the input text is parsed. The input sentence is parsed using a Minipar Parser [14] and a dependency structure is constructed from the parse tree. The Minipar parser generates a f-structure which encodes the grammatical relation of the input sentence (like subject, object and tense). This information of a sentence is represented as a set of attribute-value pairs wherein the attribute corresponds to the name of a grammatical symbol. Subsequently, the English f-structure generated is converted to ISL f-structure by applying proper transfer grammar rules using Lexical Selection and then Word order correspondence. Due to the absence of ISL written orthography, evaluation of this system is difficult. It can only be done using ISL experts.

E. Machine Translation of Indian Signs for Endocrinologist

The authors of this paper [4] have developed a translation system specific for Endocrinologists. Firstly, they have collected a lot of data for developing this translation system. The data consists of all the possible conversations that can happen between the doctor and the patient. The conversations can be the questions that the doctor asks or the suggestions given by the doctor.

The system takes the input text in plain English. The input is then parsed and tokenized. A parse tree is then formed after tokenization according to the English rules i.e. 'Subject-Verb-Object'. This text is then converted into Indian Sign Language text. Since, there are no predefined rules for the ISL text, the authors of this paper have taken the grammar rules for the ISL
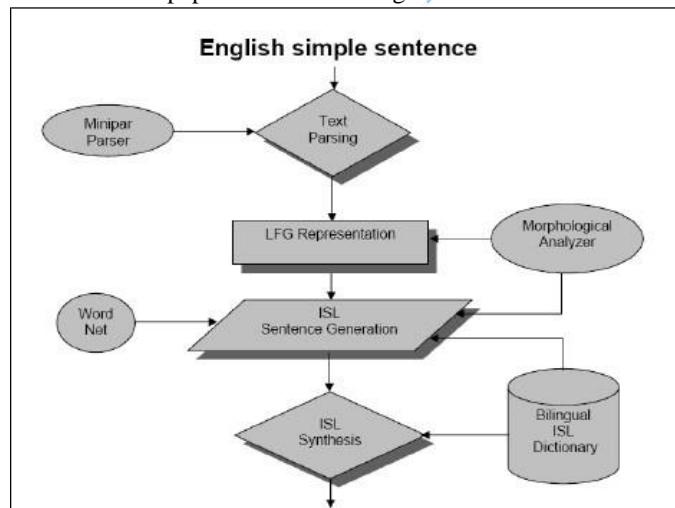

Figure 2. Architecture of the Text-to-ISL MT system

text to be 'Subject-Object-Verb'. The text obtained after this step can be mapped word by word to the signs used in the sign language.

This system like others is specific to a single field only. Also, they have not provided any analysis of the working of this system as to how efficiently this system works.

IV. CHALLENGES IN SIGN LANGUAGE

There are a lot of challenges in Sign Languages. Unlike American Sign Language (ASL) or British Sign Language (BSL), the amount of work done in India to develop Indian Sign Language is really inadequate. Many English words do not have an assigned sign. They have to be spelled in Sign Language. Also, there are differences in the dialects spoken across India. ISL has its own syntax and grammar. To transform one language into another, a detailed knowledge about both languages is needed. Machine Translation of sign language is a tough task due to lack of linguistically annotated and well documented data on ISL.

A. Survey Conducted

We conducted a survey [15] among students of age group 12-25 in various blind schools in Chandigarh to know better, the problems faced by the deaf and mute community in a real time scenario as opposed to an assumption based on individual opinions of what are the problems they are facing.

They were asked if they are aware of any such systems available. As shown in Figure 3, 90% of the students were not aware of any such existing systems and almost everyone was interested in some application which would help them in communicating with other people. When asked about the platform preferred, they were keen on the development of some kind of mobile application and that too on android (shown in Figure 4) as 95% of the students were having android phones while only 2 students were having windows phone and 1 was not having a smartphone. Currently, almost

all of them preferred using imo messenger as a means to communicate. They use it for video chatting. We also asked the students their preferred means of communication. Figure 5 shows only a few students preferred Braille over sign language. According to our survey, they also face problems in communicating using sign languages when they go to other places due to the lack of a standardized ISL.

As shown in Figure 4, almost everyone agreed to the creation of a mobile application which could help them communicating with normal people. This means that there is a desperate need of such systems. If a unified system preferably a mobile application is created for this, then the world would be such a better place for them.

V. CONCLUSION

In this paper, we have compared the existing models for the conversion of plain English/Hindi Text into the Indian Sign Language Text. Most of the systems which have been developed so far are meant to work only for a single domain. A unified system is needed for the deaf-mute community to help them interact with the normal people in their day to day lives. We propose scaling these kind of translation systems on android platform so they can be used by these people in their daily lives without any discomfort.
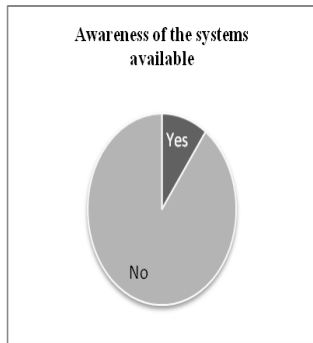
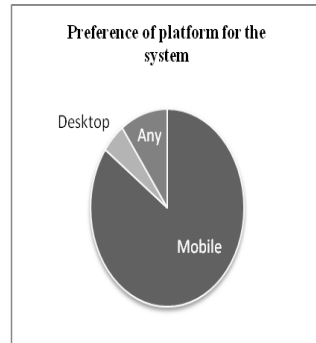Figure 3. Awareness of the systems available.

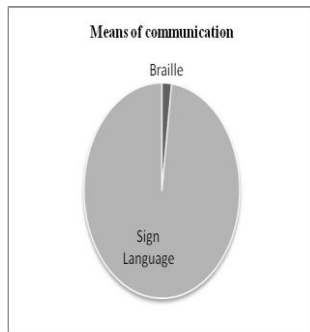Figure 4. Preference of the platform for the system

Figure 5. Means of communication

VI. REFERENCES

[1]. Kaur, Sandeep, and Maninder Singh. "Indian Sign Language animation generation system." Next Generation Computing Technologies (NGCT), 2015 1st International Conference on. IEEE, 2015.

[2]. Vij, Paras. Translator of Hindi Text to ISL and extension of ISL dictionary with WordNet. Diss. THAPAR UNIVERSITY PATIALA, 2016.

[3]. Kar, Purushottam, et al. "Ingit: Limited domain formulaic translation from hindi strings to indian sign language." ICON (2007).

[4]. Mishra, Jagriti, et al. "Machine Translation of Indian Signs for Endocrinologist."

[5]. Dasgupta, Tirthankar, Sandipan Dandpat, and Anupam Basu. "Prototype Machine Translation System From Text-To-Indian Sign." NLP for Less Privileged Languages (2008): 19.

[6]. Ali, Syed Faraz, GOURI SANKAR MISHRA, and Ashok Kumar

Sahoo. "Domain bounded English to Indian sign language translation model."International Journal of Computer Science and Informatics (2013): 41-45.

[7]. National Sample Survey Organization. Disabled persons in India. NSS 58th round (JulyDecember 2002) Report no. 485 (58/ 26/ 1). New Delhi: National Sample Survey Organization, Ministry of Statistics and Programme Implementation, Government of India, 2003.

[8]. The Deaf Community Of India. 1st ed. 2012. Web. 30 Sept. 2016.

[9]. Achraf Othman and Mohamed Jemni,"Statistical Sign Language Machine Translation: From English written Text to American sign Language Gloss", IJCSI, Vol.8 September 2011,pp.65-73

[10]. Matthew P.Huenerfauth ,"A survey of American Sign Language Natural Language Generation and Machine Translation System".

[11]. Jayant, Chandrika et al. "Vbraile: Haptic Braille Perception Using A Touch-Screen And Vibration On Mobile Phones". ASSETS'10. Orlando, Florida, USA.: ACM, 2010. 295-296.

[12]. Kumar, A. Sujith and Shahabaz Begum. "Sign Mobiles: An Android App For Specially Able People". International Journal of Advanced Research in Computer Science and Software Engineering 4.9 (2014): 208-2014.

[13]. Subhaashini, G.V.S., S. DivyaSuganya, and TharaniVimal. "Ear Hear Android Application For Specially Abled Deaf People". International Journal of Computer Science and Engineering Communications 3.3 (2016): 1108-1114.

[14]. Lin. 1998. Dependency-based evaluation of MINIPAR. In Workshop on the Evaluation of Parsing Systems, Granada, Spain "Survey For Deaf & Dumb". *Google Docs*. N.p., 2016. Web. 4 Oct. 2016.

[15]. Sangeetha, K. and L. Krishna Barathi. "GESTURE DETECTION FOR DEAF AND MUTE PEOPLE". International Journal of Development Research 4.3 (2014): 749-752.

[16]. "Hindi Wordnet". *Cfilt.iitb.ac.in*. N.p., 2016. Web. 4 Oct. 2016.

[17]. Nashat, Dalia et al. "AN ANDROID APPLICATION TO AID

[18]. UNEDUCATED DEAF-MUTE PEOPLE". International Journal of Computer Science and Mobile Applications 2.9 (2014): 1-8.

[19]. Aldein Omer, Majzoob Kamal, MohmedSirelkhtemAdrees, and Osama

[20]. Sheta. "Alternative Central Mobile Application Strategy To Deaf And Mute Education In Third World Countries". International Journal of Soft Computing and Engineering 5.3 (2015): 45-49.

[21]. Patil, Diksha Y. et al. "Android Application For Deaf People". Imperial Journal of Interdisciplinary Research 2.3 (2016): n. pag.

[22]. Saraswathi, Anitha et al. "Deaf Mute Communication Interpreter- A Review". International Journal of Applied Engineering Research 11.1 (2016): 290-296.

[23]. Koli, P.B. et al. "Image Processing Based Language Converter For Deaf And Mute People". IOSR Journal of Electronics and Communication Engineering 25-3