ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

WORD BASED STATISTICAL MACHINE TRANSLATION FROM ENGLISH TEXT TO INDIAN SIGN LANGUAGE

Gouri Sankar Mishra, Ashok Kumar Sahoo and Kiran Kumar Ravulakollu
Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University, Greater Noida, India
Email: ashoksahoo2000@yahoo.com

ABSTRACT

The objective of this work is to design a translation machine which can translate English text to Indian Sign Language glosses. A number of rule based approaches are identified in this regard, but our approach is based on statistical machine translation for ISL by using a corpus. The corpus is prepared by collecting glosses and sentences used in Indian Railways for announcement and conversation in public assistance counters. We have experimented this translation with the word based translation model. The methodologies are implemented by the tool MOSES.

Keywords: Indian sign language, statistical machine translation, word alignment, gloss, corpus.

1. INTRODUCTION

The deaf communities throughout the world have been communicating by making gestures. In the 18th century the commonly used gestures in a community were converted to sign languages and gradually these sign languages were recognized as official sign languages of those countries or, states. The sign Language used by the deaf community of India is Indian Sign Language (ISL). ISL is a natural language with its own set of gestures and grammar [1] [2] [3]. According to a survey (Zishan *et al.* 2004), 20% of deaf population of world live in India. This deaf community can't communicate with the normal people as the normal people can't understand the gestures made by the deaf and they also don't know how to make signs. Even for their basic necessities e.g. in hospitals, railway stations, banks, post offices, bus stand etc., they are struggling to express their contents. Researchers all over world are proposing and implementing various translation techniques to make the deaf community to communicate with others.

We have proposed a word based statistical machine translation technique [4] which is an emerging methodology to translate the sentences in the domain of Indian Railways announcement, enquiry and reservation. The corpora is made by collecting the sentences generally used at the windows of Indian Railway Enquiry office, the reservation office and the announcement made time to time on the platforms and the glosses from the dictionary prepared by Ramakrishna Mission Vivekananda University, Coimbatore [5]. The corpus is scientifically made by the help of MOSES Tool [18]. The process of word alignment is based on the IBM Model-1, 2, 3 along with Expectation Maximization Algorithm followed by the process of decoding. The output of the system is a sentence consisting of ISL glosses which is the translated textual representation of an ISL sentence corresponding to the given English text.

2. EXISTING MACHINE TRANSLATORS FOR SIGN LANGUAGES

The history of translations through machine starts in the sixties to translate Russian to English. Then the task became more fast and reliable with the development of computers and the computing capability. Today's translators use rules followed in grammar, statistical study of languages or, hybrid of both the methods. For statistical study of a language we require corpora of source and target languages. A number of corpuses by different agencies have been developed in various languages but the corpuses of sign languages are a rare resource which is challenging for translation of these sign languages.

The followings are some contributions on translation to Indian Sign Language. We shall also discuss about the methodologies followed for translating various sign languages in different countries.

2.1 Hindi to Indian sign language translation

INGIT [6] is the only reported tool we found for machine translation from Hindi language to ISL. The tool has been made to facilitate the deaf community of India during reservation of their tickets in a Indian Railway reservation counter. The system takes the sentences from the reservation clerk as input and displays the corresponding ISL gloss by HamNoSys. The system is based on formulaic approach which uses Fluid Construction Grammar [7]. However the techniques used in this system is not generic but can be used in a domain constrained environment. The system can't handle the structural divergence if used for other languages.

2.2 English to Indian sign language translation

Dasgupta *et al* [8] proposed a frame based approach by using transfer grammar. The proposed method finds out the syntactic structure and functional information of the inputted sentence by using a dependency parser. The output of this phase is represented in a language independent frame structure. The transfer grammar rules are applied on this frame to generate the ISL structure for the corresponding text. The architecture of the proposed model is shown below in Figure-1.

VOL. 12, NO. 2, JANUARY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

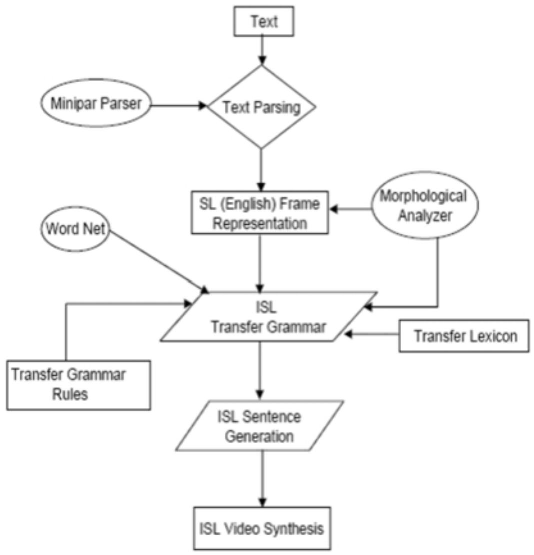

Figure-1. Architecture of frame and transfer grammar based MT for ISL.

Dasgupta *et al* [9] reported that the bilingual corpus for Indian Sign Language is not available. So the translation model follows a rule based method in which a parser known as Minipar Parser is used to parse the input sentence and a dependency structure is identified from the parse tree. This structure represents the syntactic and grammatical information of a sentence. The ISL sentence is generated from a bilingual ISL dictionary and a wordnet. From this ISL sentence the corresponding HamNoSys ISL signs are displayed. The architecture is shown in the following Figure-2.

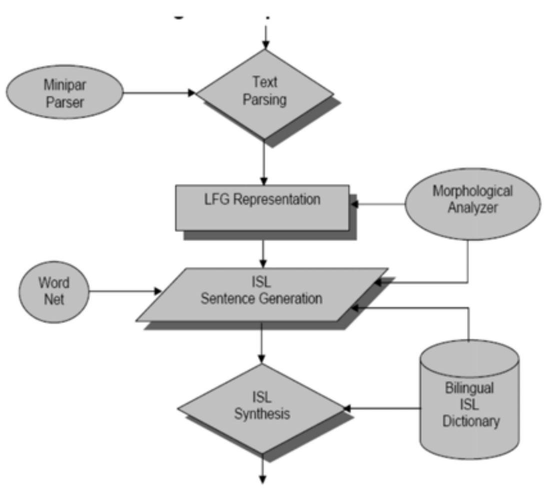

Figure-2. Prototype machine translation system.

However in both the approaches can be implemented on a very limited set of sentences. If a sentence does not follow a predefined structure then the translation of those sentences will not be feasible by the proposed methods. This means these are not generic models to be followed for a machine translation.

2.3 Other prominent sign language translations

The project ViSiCAST Translator was designed to translate English Sentences to British Sign Language (BSL) Marshall I. *et al.* [10]. The translation engine analyzes the text by a parser called CMU Link parser and by applying grammar rules it gives the output Discourse Representation structure (DRS). From this DRS Phrase Structure Rules are applied to produce the script of the sign language. The gestures are generated from this script by Gesture Sign Markup Language.

Veale *et al.* [11] the ZARDOZ system translates the English Text to American Sign Language but the system is designed in such a manner that the system can translate to British, Irish, Japanese Sign Languages also. The translation is based on rule based methodology by using Interlingua model. The system also uses Knowledge representation of Artificial Intelligence, metaphorical reasoning and blackboard system architecture. The system is domain specific as the system entirely depends on the schemata for a particular concept or, situation or, event for its syntactic and semantic analysis phase. If the source text does not follow the schemata then the system starts spell out the entire word by signing each alphabet.

TEAM [12] is project undergone at The University of Pennsylvania. The system translates the English text to American Sign Language glosses which is embedded with manual and non-manual parameters of ASL signs. The system is designed on synchronous tree adjoining grammar rules for the syntactic structure of ASL and dependency tree for English in the analysis phase. The glosses are then fetched and animated on the screen by a virtual human model as signing avatar.

R. San-Segundo *et al* [13] have evaluated three different translation methodologies from speech to Spanish Sign Language translation (SSL). The application is designed to help deaf people to apply for passport or, get similar kind of information from an officer. The three mechanisms are: Rule based methodology, statistical phrase based methodology and stochastic finite state transducers. All three mechanisms are implemented on the same corpus made from 416 sentences and 650 different words.

A. Oathman *et al*. [4] have designed a statistical machine translation system for English text to American Sign Language glosses which are then synthesized as 3D avatar signing. In the proposed work the authors implemented word based statistical translation with word alignment by the principles of IBM Model-1, 2 & 3. They have implemented the Jaro-Winkler distance mechanism for string matching after the word alignment. We found the approach is very relevant for sign language translation.

3. TRANSLATION MODEL

The architecture of the translation model is shown in the Figure-3. The source text is received by a translation module. Here the basic tasks of tokenization performed followed by a process of filtration. A specific module is performing the function of Named Entity Recognition. Then the tokens are sent to the parallel corpora for selecting the suitable glosses for each of the

VOL. 12, NO. 2, JANUARY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

tokens. Then then the process of alignment of words using IBM Model-1, 2, 3 and a learning algorithm Expectation Maximization Algorithm implemented.

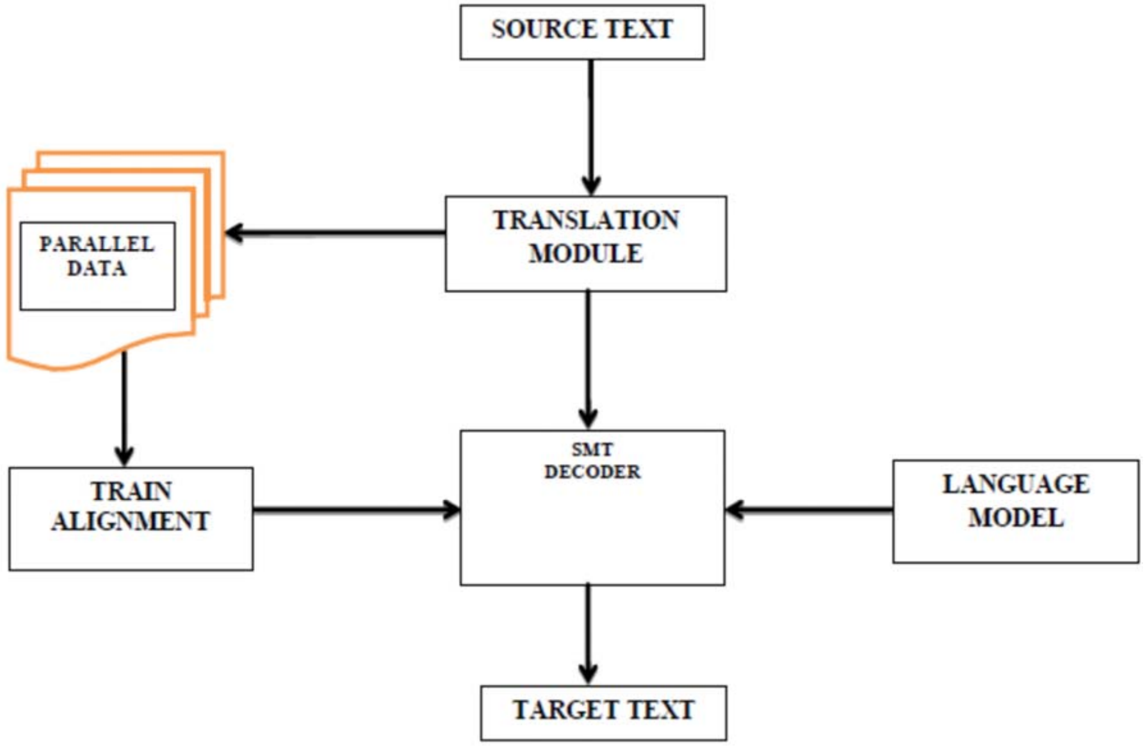

Figure-3. Architecture.

The aligned words have to undergo a phase of decoding. The decoding phase along with a language model specifies the most suitable translation for a word order. Here we have implemented 3-gram model to find out the best possible outcome of ISL glosses.

4. ISL CORPUS

As we follow the statistical machine translation we have to start the work with the translation of lexemes or, translation in word level. The first step is to find out suitable target word/s for a source word/s. A source word can have more than one meaning in different contexts as well as in different domain [15]. We can also consider the ambiguity due to homophones.

e.g.

i) Words used as Noun & Verb:

book (Verb, Reserve) : Can you please book a ticket for me?

Book (Noun, book): Can you please give me the book?

ii) Homophones:

train (A means of transport): This train is coming from New Delhi.

train (to make somebody learn skills): I had to train myself to be more assertive at work.

The challenge of ambiguity will be encountered at almost every step of translation phase as it includes the process of Natural Language Understanding (NLU) and Natural Language Generation (NLG). In fact we want to transfer the meaning of a sentence rather exchanging the words of two different languages. The new technique which works miraculously without bothering about the morphology, syntax and semantics of the source and target language is statistical machine translation [14] [16]. The reason to go for statistics is to solve the problem of ambiguity. For this purpose we require a corpus with a statistical formulation.

To translate any language to sign language statistically is very challenging due to unavailability of suitable corpus. As stated by Dasgupta *et al.* [9] they have opted for rule based system to translate English text to ISL. The corpus of a sign language is a scarce resource. We have made a corpus of 537 glosses collected from different resources [5] [17]. The glosses are the written word which represents a sign. The glosses are from general conversation and some of them are domain specific. We have collected the sentences from the domains like railway inquiry, platform announcements, railway reservation counters of Indian railways and some general conversation between two people. The total no of sentences are 326. The glosses are identified from the repository of signs of Indian sign language [17] and manually assigned to a set of English words or, a set of phrases.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

4.1 Word alignment

The sentences of the source language have to be tokenized and the words from the sentences have to be identified. The corresponding gloss for each word has to be found out from the dictionary which keeps the most likely gloss for the input word. The gloss will represent the target word of a source word or, group of source words. There may be more than one gloss for a source text also. Here we have to keep in mind that each English word does not require a gloss. For example:

Sign : NAYA: ABHI: BACHA: ISKU: L SIKHA: NA:
ISL Gloss: new now child school teach
English : The new (thing is that) children are now taught (sign language) at school.'

Sign : PAHLE SIKHA: NA: MUSKIL.
ISL Gloss: before teach difficult
English : '(Signs) were not taught before; (it was too) difficult.'

Sign : ABHI: SURU: BACHA: SIKHA: NA:
ISL Gloss: now begin child teach
English : 'Now they have started to teach the children.'

Figure-4. Examples of English to ISL gloss translation.

From the above examples [17] it can be observed that the meaning of the sentence has to be transferred. Not all the words are required for translation.

4.2 IBM model-1

In the theoretical model we assume the translation of English word 'e' to the gloss of ISL ''g' as a conditional probability function t(g|e). This function is applied on the corpus to find out the gloss of ISL. As shown in figure-5 the target ISL glosses have been found. Up to this step we have only found the lexemes to be used in the target sentence. This is a direct word to word translation mechanism.

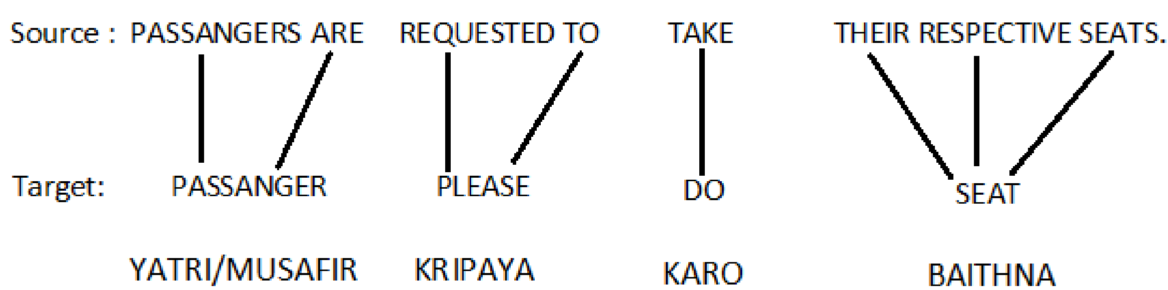

Figure-5. Mapping of words.

The word order in the target text may not be a valid sentence as per the syntax and semantics are concerned. In order to make a valid sentence from these lexemes the arrangement of lexemes is required. The alignment of words can be formulated with a function from a: i→j, where i indicates the position of English word and j represents the position of respective gloss for i. Here in our example the alignment function is represented as:

a: { 1→1, 2→1, 3→2, 4→2, 5→4, 6→3, 7→3, 8→3 }

The alignment of the source English words to the target glosses are shown in Figure-6.

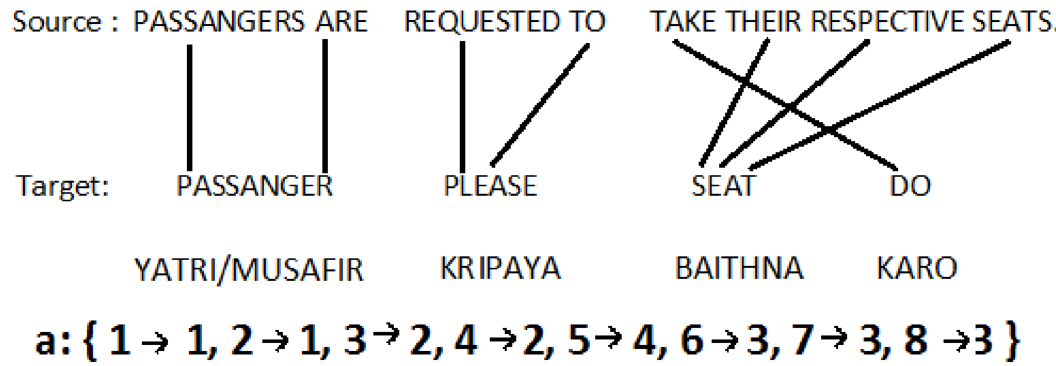

Figure-6. Word alignment in IBM model-1.

In the above procedure we have followed the IBM Model1 which can be defined as a joint probability function for the English text e = { e_1, e_2, …., e_{le}} to a gloss of Indian Sign Language sentence g = { g_1, g_2, …., g_{lg}} with an alignment function a from g_j to e_f as follows:

$$p(e,a|f) = \frac{\epsilon}{(l_f+1)^{l_e}} \prod_{j=1}^{l_e} t(e_j|f_{a(j)})$$

So, according to IBM Model-1 the probability of each of the gloss to be selected is equal and is 0.125. In IBM Model-2 this probability is not same.

4.3 IBM model-2

In IBM Model-2 [15] the arrangement probability of glosses depends on their positions. As the words in a sentence occupies a specific position according to the syntax. But instead of following the grammar of the target language we shall consider the conditional probability of a word in a sentence in respect to other words. Therefore, the probability distribution of alignment function in IBM Model-2 is a (j|f, l_e,l_g) where the length of source word is l_e and the length of target word is l_g.

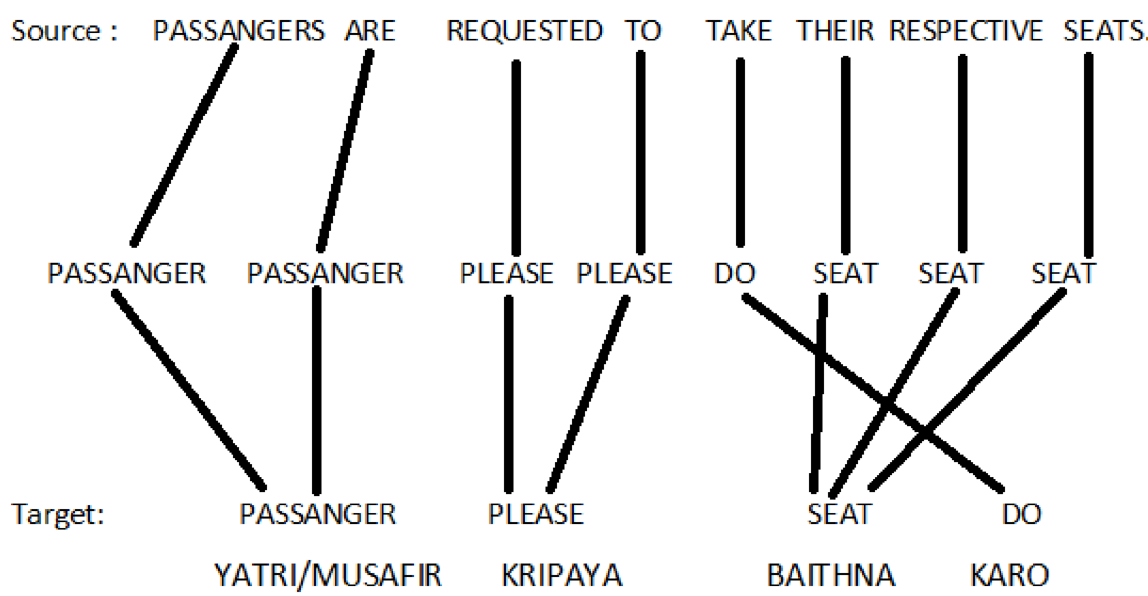

Figure-7. Word alignment IBM Model-2.

In IBM Model-2 there will be two steps of translation. The first step is exactly same as IBM Model-1 and the process of translation probability t(g|e) is implemented. Then in the second step the alignment wiil be done.

e.g. The translation of REQUEST to PLEASE has a lexical translation probability of: t(PLEASE|REQUEST). In the second step the alignment probability will be a (2|3, 4, 5).

VOL. 12, NO. 2, JANUARY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

If we combine both the steps mathematically then the combined formula will be:

$$p(e,a|f) = \epsilon \prod_{j=1}^{l_e} t(e_j|f_{a(j)})a(a(j)|j,l_e,l_f)$$

The IBM Model-2 do not consider the null target words. In Indian sign Language we come across with the target words which are not mapped from the source words. In IBM Model-3 we shall see the example and solve the problem of alignment of null target words.

4.4 IBM model-3

Let's consider the following example shown in Figure-8. The words of source text are translated to the ISL gloss but one of the glosses has not been aligned with any of the source text.

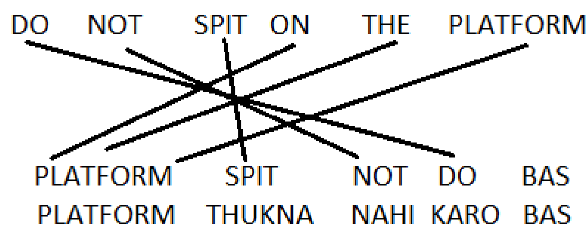

Figure-8. Gloss without alignment.

In Indian Sign Language some functional words like BAS, HO_GAYA, KARO etc. may not be aligned with any source word. Since we are translating the contents of a sentence, we have to include the glosses which are necessary to convey the original thought. In the previous examples, the word TAKE is aligned to DO, which is explicitly done in our corpus. If we consider in a corpus this is not done then it may happen that "TAKE YOUR SEAT" will be aligned to SEAT i.e. BAITH: NA and DO i.e. KARO will be a functional sign to be added to the sentence without alignments.

For the above situations the non-aligned word of the target sentence will be aligned with a null source text word. The probability of generating a null token is [15]:

$$p(\emptyset_0) = \binom{l_e - \emptyset_0}{\emptyset_0} p_1^{\emptyset_0} p_0^{l_e - 2\emptyset_0}$$

It means that we have to estimate our model from an incomplete data. We have used Expectation Maximization Algorithm to train all the IBM Models in alternating steps of learning. This iterative method of training a model overcomes the problem of incomplete data.

5. RESULTS AND WORD ALIGNMENT MATRIX

For evaluation we have trained a n-gram model on a small corpora shown in the Table-1.

Table-1. Statistics of corpora.

Language	Sentences	Tokens
English	**326**	**739**
Indian Sign Language	326	537
n-gram 1 = 511	n-gram-2 = 1354	n-gram-3= 276

In a n-gram model the n-1 number of places are considered for a word to be placed in a particular position. For the incoming word at the position say k will be decided by a probability function of the language model as $P_{LM}(W_k|W_{k-(n-1)}, W_{k-(n-2)}\ldots W_{k-1})$ for a word W_k.

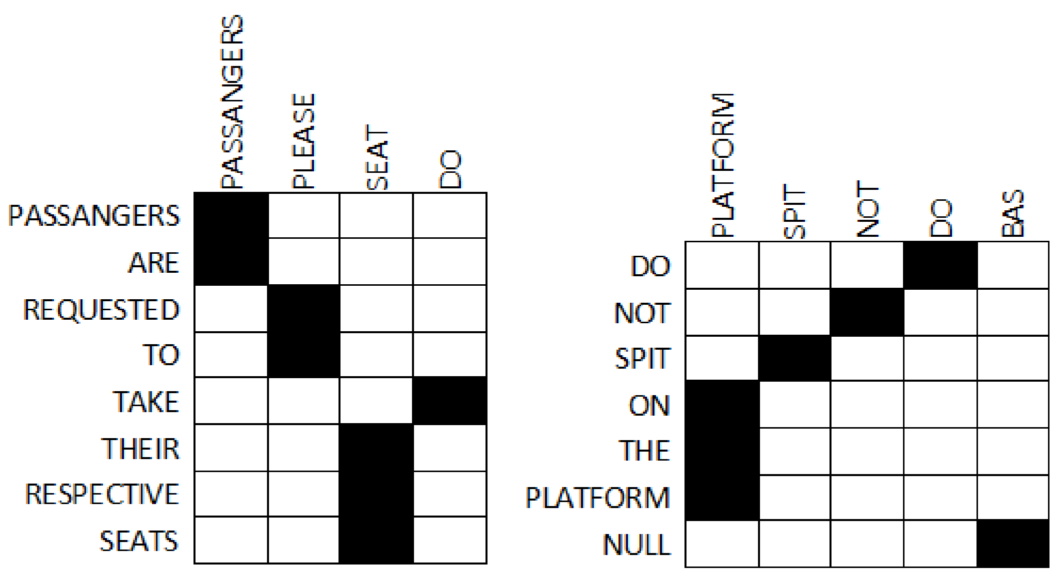

Figure-9. Word alignment matrices.

ARPN Journal of Engineering and Applied Sciences

The resultant of a word alignment can be represented in the form of a matrix as shown in the Figure-9. The Y-axis shows the text that has been given as input to the translation machine and the aligned words are arranged in X-axis. From the word alignments of both the matrices it can be observed that one to one alignment is not mandatory. The words may have multiple alignment points e.g. ‘ON’, ‘THE’ and ‘PLATFORM’ are aligned to ‘PLATFORM’ gloss. Some words may have NULL alignment points, e.g. BAS is aligned to NULL.

6. CONCLUSIONS

So far as statistical machine translation is concerned we have reported the first communication for Indian Sign Language with a small corpus containing sentences from various conversation of public interfacing in Indian Railways. In the first step we have contributed to lexical translation model in which words of source sentences are translated to the words of target language and the order of these words are changed to make a valid target sentence. The dependency on the rules and grammar of both the languages are ignored. The work can also be improved by introducing phrase based statistical machine translation as the word based translations can be implemented successfully on the sentences with less number of words. For larger sentences the method may not work with better accuracy. So the work can be improved by implementing phrase based machine translation technique.

ACKNOWLEDGEMENT

We are thankful to Late Prof. Pervez Ahmed for his inspiration and guidance for the work.

REFERENCES

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

[2] Zeshan U. 2003. Indo-Pakistani Sign Language Grammar: A Typological Outline. Sign Language Studies. 3(2): 157-212.

[3] Zeshan U., Madan M. Vasishta, Sethna M. 2004. Implementation of Indian sign language in educational settings. 15(2), Asia Pacific Disability Rehabilitation Journal. pp. 15-35.

[4] Othman A., Jemni M. 2011. Statistical Sign Language Machine Translation: from English written text to American Sign Language Gloss. 8(Issue 5, No 3): International Journal of Computer Science Issues, pp. 65-73.

[5] http://indiansignlanguage.org/dictionary/ (Accessed on 15 Jul 2016).

[6] 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 (ICON), Hyderabad, India.

[7] Steels L., Beule J.D. 2006. Unify and Merge in Fluid Construction Grammar. In: Lyon, C., Nehaniv, L. and A. Cangelosi (eds.), Emergence and Evolution of Linguistic Communication, Lecture Notes in Computer Science, Springe-Verlag: Berlin.

[8] Tirthankar Dasgupta and Anupam Basu. 2009. Indian Sign Language Machine Translation System for the Deaf People of India, Alternative and Augmentative Communication and Assistive Technologies (AACAT). pp. 44-45, Mumbai, India.

[9] Dasgupta T., Basu A. 2008. Prototype machine translation system from text-to-Indian sign language. Proceedings of the 13th international conference on intelligent user interfaces, Gran Canaria, Spain. pp. 313-316.

[10] 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 Language Processing. pp. 154-159, Tzigov Chark, Bulgaria.

[11] T. Veale, A. Conway, B. Collins. 1998. The challenges of cross-modal translation: English to sign language translation in the ZARDOZ system. in Machine Translation. 13. 81-106.

[12] C. J. Wideman and E. M. Sims. 1998. Signing Avatars. Technology and Persons with Disabilities Conference.

[13] R. San-Segundo, A. Pérez, D. Ortiz, L. F. D’Haro, M. I. Torres, F. Casacuberta. 2007. Evaluation of Alternatives on Speech to Sign Language Translation. Proc. of Interspeech 2007, pp. 2529-2532. Antwerp, Belgium.

[14] P. Koehn, F. Och, D. Marcu. 2003. Statistical phrase-based translation. NAACL'03 Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. pp. 48-54.

[15] Book: P. Koehn, Statistical Machine Translation, Cambridge University Press, 2009.

[16] P. Brown, V. Pietra, S. Pietra, R. Mercer. 1993. The mathematics of statistical machine translation: parameter estimation. Computational Linguistics - Special issue on using large corpora: II, Vol. 19, pp. 263-311.

[17] Book: Ulrike Zeshan. SIGN language in Indo-Pakistan a description of a signed language, La Trobe University.

[18] Philipp Koehn. Moses: Statistical Machine Translation System User Manual and Code Guide. University of Edinburgh.