

# Communications between Deaf and Hearing Children Using Statistical Machine Translation

Mahdiyeh Alimohammadi\*, Moteza Zahedi

College of Computer Engineering, Shahrood University, Iran

Copyright © 2015 Horizon Research Publishing All rights reserved.

**Abstract** Communication with hearing people society is an important problem for deaf people. They have not learned the valid rules of spoken language that hearing people use them. We therefore prepare an efficient corpus and apply it to Moses Machine Translation to simplify these communications. We choose communications between children because they use e-communications more than adults. All systems that automatically process sign language corpus rely on appropriate data. So our corpus with a limited set of words and with specific subject is the first Persian corpus containing Persian language, PL, and Persian sign language, PSL, based on the domain of children's conversations. At the first step raw data are pre-processed which provides necessary information for translation. These data are statistic information extracted of sentences. After getting important data from initial sentences, the corpus is applied for training of Moses machine translation. In addition to the main goal of this system, we can educate deaf people the valid Persian grammar that is a problem for deaf people in school and society. In this paper, we compare our results with the results taken from the Moses decoder in other spoken languages that indicate our purpose is applicable in the real world.

**Keywords** Persian Sign Language, Children Communications, Statistical Machine Translation, Sign Language

---

## 1. Introduction

The Statistical Machine Translation (SMT) is a challenge for the field of natural language processing. Nevertheless, the systems capable of such processing are useful in many applications such as helping deaf people that have daily communication challenges in a hearing environment. Traxler 2000 [1], shows that the majority of the deaf society has weak reading skills and they cannot properly communicate with hearing people because they do not use the rules of spoken language that hearing people use them. So this

problem makes it more difficult to help these people for automatic aids.

Moreover, deaf children can learn how to communicate with hearing people. Childhood is an appropriate period of time for learning true communication methods. In the community we have been observed that deaf people that are able to communicate with others, have a better understanding of their environment and also have more ability to understand and learn. Therefore, automatic translation and recognition systems are developed in this area. However, for the modern statistical machine translation (SMT) it is necessary to present a set of appropriate data. But unfortunately, the most of currently available corpora are too general or too small in a specific domain for training the automatic recognition and statistical machine translation systems of sign language. In Persian language there is no appropriate corpus for this purpose. We therefore in this paper have prepared a sign language corpus in Persian language and Persian sign language for a specific domain of children's conversations in the application of basic Statistical Machine Translation (SMT) and sign language analysis. To achieve this purpose, after pre-processing this corpus (Persian Sign Language Corpus or PSLC) is ready for Moses machine translation that is a good translator for our purpose.

Our paper is outlined in the following. After a description of related works in section 2, we explain our corpus and prepare it for pre-process and the main goal, translating, in section 3. In the section 4 we describe our selected machine translation and purpose of our choice, pre-processes, necessary tools for this purpose, the operation mode of machine translation and also the necessary steps for making the correct translation. After the introduction of evaluation metrics, we present our results in section 5 and show this work in comparison with other spoken translations can be applied in the real world in this section. Finally, we express conclusion and future works in section 6 for the future.

## 2. Related Works

Some researchers worked on sign language corpora, but

most of them focused on linguistic aspects rather than natural language processing. Some of these works are:

The European Cultural Heritage Online organization (ECHO)<sup>1</sup> as stated in [8], published corpora in Swedish sign language, British sign language and the sign language of the Netherlands. These data contain children’s fables and poetry each signed by a single signer. However, they have a large vocabulary which makes automatic learning difficult. Though, in a related work, Morrissey and Way, 2005 [2] applied example based methods for automatic translation based on one of the ECHO corpora. However, their results imply that their system is only robust for sentences already seen in training, but has problems with unseen words and phrase combinations.

The American Sign Language Linguistic Research group at Boston university created a set of videos in American Sign Language which is partly available on their website<sup>2</sup> and described in (Neidle et al., 2000) [3]. All videos are annotated and recorded in three different perspectives. (Zahedi et al., 2005) [4], as stated in [8], published results on sign language recognition for this corpus. The corpus focuses on linguistic topics.

(Heßmann, 2001) [5], as stated in [8], published a corpus based on interviews in DGS with several thousand sentences. This is publicly available on the ECHO website too. While this corpus is quite large, its domain is too broad, making automatic learning difficult.

The Ireland corpus developed at the center of deaf Studies, Dublin (Leeson et al., 2006) [6]. This corpus contains recorded videos of 40 Irish signers aged between 18 and 65 that are collected over three years. They tell the personal selected narrative, a story that is for children with ‘The Frog’ name and sign elicited sentences.

(Y.-H. Chiu and Cheng, 2007) [7] perform SMT experiments on a corpus of 1983 sentences for a pair of languages Chinese and Taiwanese sign language.

Phoenix corpus, a corpus of 2468 sentences For the domain of weather reports in German and DGS that was presented by (Bungeroth et al., 2006) [8]. It is particularly used for SMT and sign language recognition. DGS translation is provided by an interpreter in German television channel Phoenix part of the weather report of daily news.

The RWTH-BOSTON-104 Database (Dreuw et al., 2007) [9], This corpus that is mainly used for automatic sign language recognition contains 201 sentences and 104 words in American Sign Language with its English annotations. The sentences are signed by two females and one male.

ATIS In (Bungerot et al., 2008) [11], a corpus of 595 (small size) sentences in five languages English, German, Irish Sign Language, German Sign Language and South African Sign Language that is in the domain of Air Travel Information System (ATIS). With being a multilingual dataset, it is possible to directly translate between sign languages.

Our corpus is useful for automatic recognition and

statistical machine translation of sign language. Similar to Phoenix database, we present the PSLC dataset based on the domain of children’s conversations that is useful for the scope of recognition and SMT efficiently. We have created a lot of sentences with a limited set of words because restricted word variance increases efficiency of the systematic training.

### 3. Corpus Setup

All sentences of the original PSLC dataset are given as a text file in Persian to form the transcription of spoken sentences with the Persian sign language. Table 1 gives several examples that are selected from the Persian sign language Corpus. Totally, 2020 sentences made of 150 words are translated by three deaf native to PSL, one man and two women. For a better view, we translate both groups into English in Table 1.

**Table 1.** PSLC Data Example

Original sentences in English	Converted sentences to sign language	Original Persian sentences	Converted sentences to PSL
Maryam also sent to me.	Maryam send me.	مریم برای من هم فرستاد.	مریم من فرست.
I solve the math practices.	I solve the math practice.	من تمرینهای ریاضی را حل کردم.	من تمرین ریاضی حل کرد.
Davoud was at our home today.	Davoud our home today.	امروز داوود خانه ی ما بود.	داوود امروز خانه ما.
I lost the questions. It is an interesting music, is not it?	I question was lost. Interesting music?	من سوالات را گم کردم. آهنگ جالبی است نه؟	من سوال گم شد. آهنگ جالب؟

In our current corpus, data are in XML format. Each sentence in XML file includes these fields for using in machine translation: Persian Language, Persian Sign Language, the path of files and an English version of the Persian language sentences. Each record is identified by a unique identifier.

**Table 2.** Corpus Overview

Statistics	Training	Development	Test
No.sentences	1620	250	250
No.words	9856	1342	1297
No.single words	150	76	82

For using our corpus in SMT systems we should split it into a training set, a development set and a test set. The training set is used for the learning of the system, where repeated phrases are memorized, development set helps to optimize the parameters of the system and test set is used for evaluation. Table 2 shows a detailed breakdown of the sets

<sup>1</sup> <http://www.let.kun.nl/sign-lang/echo/>

<sup>2</sup> <http://www.bu.edu/aslirp/>

that is used in our MT application.

## 4. Machine Translation

The first step after data gathering is the preparation of the parallel corpus data. It should be converted to XML format and then be tokenized, lowercased and sentences which would be too long to handle (and their correspondences in the other language) have to be removed from the corpus. For translating Persian sign language to Persian language and vice versa, we need a machine translation. So once the corpus data is prepared, the actual training process can be started with machine translation. We choose Moses machine translation. The main reasons for our choice are as follows:

**Beam-search:** an efficient heuristic search algorithm that finds the highest probability or the best translation among top best translations that are selected according to sentence structures from the exponential number of choices.

**Phrase-based:** it is translating sequences of words that may be with different lengths by statistical methods. This form of translation reduces the restrictions of word-based translation.

**Factored:** words with factored representation. This model augments phrase-based SMT with layered dependencies, and with a joint model that extends the phrase translation table with micro-tags, i.e.

Decoding of confusion networks, enabling easy integration with ambiguous upstream tools, such as automatic speech recognizers or morphological analyzers [24].

Novel factored translation models, which enable the integration of linguistic and other information at the word level at many stages of the translation process [24].

Support for large language model. The Moses system implements language model so that is more efficient than the canonical SRILM (Stolcke 2002) implementation used in the most systems. So a minimum loading time is required [15].

With implementation of the Moses decoder<sup>3</sup> which was developed by a large group led by Philipp Koehn and Hieu Hoang [Koehn et al., 2007] [15], a fully compatible MT system became available that was open for modifications. For implementation of our goal, we have used the Moses decoder. This decoder is designed within a strict modular and object-oriented framework for easy maintainability and extensibility [17]. For decoding and translation, we have used some necessary tools that are executed before decoding process in Linux operating system [12][13], for creation of language and translation models that are fed to the global search for translation. Language models are created using the SRILM toolkit [19], word alignment during the training is done with GIZA++ [21] and we used mkcls [22] that is a tool for training word classes using a maximum-likelihood criterion.

### 4.1. Requirement Processes

The SRILM is a toolkit [Stolcke, 2002] for creating and applying various statistical language models (LMs) that can be used in statistical machine translation, speech recognition, statistical tagging and segmentation. In machine translation, SRILM language models are currently available in the Moses MT system. The SRILM toolkit is designed and implemented by Andreas Stolcke. The most important part of this toolkit is the n-gram-count tool which counts n-grams and estimates language models that are used in machine translation. The main three Functionalities of this toolkit briefly are: generating the n-gram count file from the corpus, training the language model from the n-gram count file, calculating the test data perplexity using the trained language model.

The next step is alignment between a pair of strings as a measure to find for each word in PLS string the corresponding word in PL string. For this purpose, Giza is a training program that learns statistical translation models from bilingual corpora. GIZA++ [Och and Ney, 2003], an extension of GIZA, is a tool to compute word alignments between two sentences aligned corpora, which is useful for both sentence alignment and phrase table generation. Some translation models introduced by IBM scientists in early 1990s are used in Giza++, included IBM model 1-5. After constructing translation models, with IBM models, the output is alignment file. The first line of this file is a label that can be used as a caption for alignment visualization tools. It contains information about the sentence sequential number in the training corpus, sentence lengths, and alignment probability. The second line is the target sentence; the third line is the source sentence. Each token in the source sentence is followed by a set of zero or more numbers. These numbers represent the positions of the target words to which this source word is connected, according to the alignment. Before data processing, the parallel corpus should be converted into a suitable format for GIZA++ toolkit. Then, two vocabulary files are generated and the parallel corpus converted into a numeric format. Alignment models generated depending on word classes of mkcls.

mkcls [Och, 1999] is a tool for training word classes using a maximum-likelihood criterion. GIZA++ requires words to be placed into word classes. The resulting word classes are especially suited for language models or statistical translation models. For using the IBM-Model4 or HMM Model in GIZA++, we need to word class files. We use the mkcls to reach this. Both GIZA++ and mkcls were designed and implemented by Franz Josef Och. Both of these tools will be called by Moses training scripts.

### 4.2. Machine Translation Setup

Sentence decoding is done with the Moses decoder. As mentioned before, Moses is a statistical machine translation system that trains translation models for any given language pair in a parallel corpus. What we need is a set of translated texts or parallel corpus. The Moses decoder works using a

<sup>3</sup> . <http://www.statmt.org/moses/>

beam search [Koehn, 2004a] algorithm that determines the best translation for a given input. The decoder is the core component of Moses. To minimize the learning curve in many of researches, the decoder was developed as a drop-in replacement for Pharaoh [20], the popular phrase-based decoder [15] that if you present a translation model and language model, it translates the source language into the target language. During the process, phrase tables are built by Giza++ word alignments and then the best translation for new input is built using phrase table plus SRILM language model. Moses MT is phrase-based [Koehn et al., 2003] and allows words to have a factored representation. All of these processes as compact boxes are shown in Fig. 1. This is a theoretical description of the translation process. This theory is described in detail in [10].

Totally in statistical machine translation, we are given a source language sentence such as  $f_1^J = f_1 \dots f_j \dots f_J$ , which should be translated to a target language sentence  $e_1^I = e_1 \dots e_i \dots e_I$ . For this claim we should choose the target sentence with the highest probability. We assume that  $\Pr(e_1^I)$  is the "Language model" which is estimated using monolingual corpora and contains a large number of words and their probability of occurrence.  $\Pr(f_1^J | e_1^I)$  is the "Translation model", which is estimated by bilingual corpora and is the probability of translating source 'f' from target 'e'. The translation process is stated in Equation.1 and is outlined in Fig.1.

$$\hat{e}_1^I = \arg \max_{e_i^I} p(e_1^I | f_1^J) = \arg \max_{e_i^I} p(e_1^I) p(f_1^J | e_1^I) \quad (1)$$

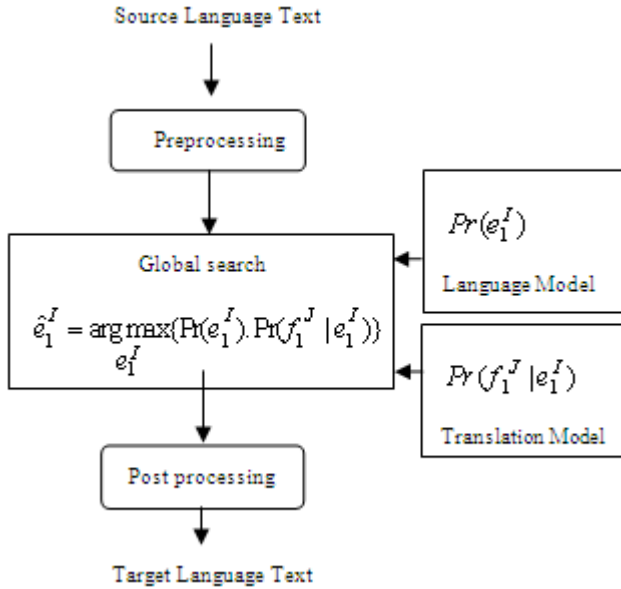


Figure 1. Translating process from language "f" to language "e"

This approach in Fig.1 is referred to as source-channel approach for statistical MT. Sometimes, it is also referred as the 'fundamental equation of statistical MT' (Brown et al.,

1993) [10]. For decoding we have a language model, a translation model and a new sentence  $f$ . We find translation  $e$  maximizing  $\Pr(e_1^I) \times \Pr(f_1^J | e_1^I)$ . Language model can be calculated with n-grams in SRILM toolkit. For calculating translation model, as stated in [25], using the equation  $\Pr(f | e) = \text{count}(f, e) / \text{count}(e)$  (for simplifying,  $\Pr(f | e)$  instead of  $\Pr(f_1^J | e_1^I)$ ), is not possible. Because the sentences are novel and we never have enough data to find it for all of the sentences. So we decompose the sentences into smaller chunks, like in language modeling and introduce another variable  $a$ , to represent alignments between the individual words in a sentence pair:

$$p(f | e) \equiv \sum_{\alpha} p(\alpha, f | e) \quad (2)$$

So we can calculate translation probabilities using these alignment probabilities. Now we need to define  $p(\alpha, f | e)$  as Equation.3.

$$p(\alpha, f | e) = \prod_{i=1}^m t(f_j | e_i) \quad (3)$$

Where  $e_i$  is the target language word(s) corresponding to the source word  $f_j$  and  $t(f_j | e_i)$  is conditional probability of the words being aligned. Relative frequency estimates can be used to estimate  $t(f_j | e_i)$ , but the most problem is we do not have word-aligned data, only sentence-aligned. There is an elegant mathematical solution to this problem; the EM algorithm can be used here. These processes are running in Moses decoder. To translate an observed source sentence  $f$ , we seek the target sentence  $e$  which maximizes the product of those two terms (language model and translation model) which is called decoding.

## 5. System Evaluation

### 5.1. Evaluation Metrics

After translation, we should evaluate the results that are necessary for ranking systems and evaluate incremental changes. We cannot use human evaluation because it is very time-consuming and it is not re-usable. Therefore, we use automatic evaluation. The basic reasons for automatic evaluation are stated in [25] and are: (1) no cost evaluation of incremental changes, (2) ability to rank systems, (3) ability to identify which sentences are doing poorly on, (4) categorize errors, and (5) correlation with human judgments and Interpretability of the score. Some metrics for evaluation are WER, PER, NIST and BLEU [Papineni et al., 2002] [16] that we used in our work and describe them as follows.

Word Error Rate (WER): The word error rate as its calculation is described in (Hunt, 1990) [18], commonly used for evaluating automatic speech recognition or machine translation systems. This criterion is derived from the

Levenshtein-distance, or edit-distance. The edit distance between two strings is the minimum number (or weighted sum) of insertions, deletions and substitutions for transforming source string into another [25]. The WER is the edit distance between a sequence of words as reference sequence and its translation sequence that is normalized according to the length of the reference sequence. This normalization is done for making possibility to compare different systems with different tasks. For this purpose we define: N as total words in the reference sequence, S is the number of substituted words in the automatic translation, D is the number of words of the reference sequence that are deleted from the automatic translation and I is the number of words that are inserted in the automatic translation and are not in the reference sequence. The Word Error Rate (WER) is calculated as (4):

$$WER = \frac{S + D + I}{N} \quad (4)$$

Position independent word Error Rate (PER): A related metric is the Position independent word error rate (PER) [14]. In this evaluation criterion, re-ordering of words and sequences of words is allowed between a translated transcription and a reference translation. In Word Error Rate, order of words in sequence is important and this is a problem of WER. May be there are two sequences with different order of words, but both of them are acceptable so that the WER criterion, alone, could be misleading and for overcoming to this problem, we introduce an additional measure with the name of Position independent word Error Rate (PER). This criterion compares the words of two sentences without taking the word order into account. The PER is guaranteed to be less than or equal to WER.

BLEU score (bilingual evaluation understudy): This criterion is based on the precision and it is a measure of correspondence between one translated sentence and a set of reference sentences by calculating geometric mean of n-grams precisions in both of the sentences. With more references, possibility for being good references will increase and then criterion precision will be high. This geometric-mean is multiplied by a factor  $BP(.)$  that penalizes short sentences:

$$BLEU = BP(.) \cdot \exp\left(\sum_{n=1}^N \frac{\log p_n}{N}\right) \quad (5)$$

Where  $p_n$  denotes the precision of n-grams in the hypothesis translation. The goal of this metric is distinguishing which system has better quality (correlation with human judgments). A higher score represents a more accurate translation with this criterion. An appropriate tool for this is the NIST BLEU scoring.

NIST score: This criterion is based on the BLEU metric, but calculates weighted precision of n-grams between a candidate and a set of reference translations instead of calculating precision of n-grams with equal weights. In this criterion n-grams with less likely to occurring will receive lower weight than n-grams with more likely to occurring.

This measure is multiplied by a factor  $BP(.)$  as (6), that penalizes short sentences.  $w_n$  is weighted precision of n-grams.

$$NIST = BP(.) \cdot \sum_{n=1}^N w_n \quad (6)$$

Both, NIST and BLEU are two accuracy measures and so the larger values reflect better translation quality.

## 5.2. Experimental Results

We used n=3 and so Three-gram in our experiments. For evaluation, we want to train a translation model for PSL to PL. The training process takes a lot of time and memory, about 25 minutes in our system with one core CPU and 512MB RAM to complete. After running translation process, we got efficient results with these evaluation metrics in section 5-1. Table 3 shows evaluation results for our corpus with Moses decoder. We compare this system with Czech to English and German to English translations of newswire text, a scenario in which SMT usually excels [26], by using Moses decoder and Cunei machine translation [26][27]. Translation results these languages are demonstrated in table 4. As is compared, our results represent translation technique on this corpus is applicable in real world for helping deaf people to communicate with hearing people society in web environment.

**Table 3.** Our Evaluation Result Of Translation

PER	WER	NIST	%BLEU
19.16%	45.69%	5.36	30.05

**Table 4.** Test Set Performance In [25][26] For Comparison

Language pair	Moses		Cunei	
	%BLEU	NIST	%BLEU	NIST
Czech-English	20.41	6.4794	21.16	6.6221
German-English	20.60	6.5102	21.11	6.5639

## 6. Conclusions and Future Work

In this paper we have presented the PSLC for both of Persian language and Persian Sign Language (PSL) of the domain of children conversation, which is suitable for sign language recognition and sign language translation. It is the first corpus in Persian sign language that has sufficient data in a limited scope. Then, we have prepared our corpus in an efficient machine translation that will be used in online communication between deaf and hearing children. We got the results that this technique is sufficient for our purpose, easy communication between deaf children with hearing people. Our program for future is to prepare the video corpus completely and use them in video chatting in addition to corpus translation. For this purpose, as we briefly described, we prepared 200 video of three deaf people with a webcam. Since webcam is more common for online chatting. Further development of this work will be used for a full

communication between deaf and hearing children.

---

## REFERENCES

- [1] C. B. Traxler, "The Stanford Achievement Test: National norming and performance standards for deaf and hard-of-hearing students," *Journal of deaf studies and deaf education*, Vol. 5, No. 4, pp. 337-348, 2000.
- [2] S. Morrissey, & A. Way, "an example-based approach to translating sign language," In *Second workshop on example-based machine translation, MT summit X workshop*, Phuket, Thailand, 2005, pp. 109-116.
- [3] C. Neidle, J. Kegl, D. MacLaughlin, B. Bahan, & R G. Lee, "The syntax of American Sign Language: Functional categories and hierarchical structure," Cambridge, MA: MIT Press, 2000.
- [4] M. Zahedi, D. Keysers & H. Ney, "Appearance-Based Recognition of Words in American Sign Language," In *IbPRIA 2005, 2nd Iberian Conference on Pattern Recognition and Image Analysis*, Estoril, Portugal, Volume LNCS 3522 of *Lecture Notes in Pattern Recognition and Image Analysis*, June 2005, pp. 511-519.
- [5] J. Heßmann, "Gehörlos so! Materialien zur Gebärdensprache. Band 1: Grundlagen und Gebärdenverzeichnis," Hamburg, Germany: Signum Verlag, 2001.
- [6] L. Leeson, J. Saeed, C. Leonard, A. Macduff, & D. Byrne-Dunne, "Moving Heads and Moving Hands: Developing a Digital Corpus of Irish Sign Language," The 'Signs of Ireland' Corpus Development Project, In *Proceedings of Information Technology and Telecommunications*, Carlow, Ireland, 2006.
- [7] Y. H. Chiu, C. H. Wu, H. Y. Su, & C. J. Cheng, "Joint optimization of word alignment and epenthesis generation for Chinese to Taiwanese sign synthesis," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, Vol. 29, No. 1, pp. 28-39, 2007.
- [8] J. Bungeroth, D. Stein, P. Dreuw, M. Zahedi & H. Ney, "A German sign language corpus of the domain weather report," In *Fifth International Conference on Language Resources and Evaluation*, Genoa, Italy, 2006, pp. 2000-2003.
- [9] P. Dreuw, D. Rybach, T. Deselaers, M. Zahedi & H. Ney, "Speech recognition techniques for a sign language recognition system," In *Interspeech 2007 - Eurospeech*, Antwerp, Belgium, 2007, pp. 2513-2516.
- [10] P. F. Brown, V. J. D. Pietra, S. A. D. Pietra, & R. L. Mercer, "The mathematics of statistical machine translation: Parameter estimation," *Computational linguistics*, Vol. 19, No. 2, pp. 263-311, 1993.
- [11] J. Bungerot, D. Stein, P. Dreuw, H. Ney, S. Morrissey, A. Way & L. van Zijl, "The ATIS sign language corpus, In 6th International Conference on Language Resources and Evaluation," Marrakech, Morocco, 2008, pp. 4
- [12] C. Federmann, "Very large language models for machine translation," Master's thesis, Universität des Saarlandes, Saarbrücken, Germany, 2007.
- [13] X. Wang, "Evaluation of two word alignment systems," Doctoral dissertation, Linköping, 2004.
- [14] C. Tillmann, S. Vogel, H. Ney, A. Zubiaga & H. Sawaf, "Accelerated DP based search for statistical translation," In *Proceedings of the 5th European Conference on Speech Communication and Technology*, Rhodes, Greece, 1997, pp. 2667-2670.
- [15] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi & E. Herbst, "Moses: Open source toolkit for statistical machine translation," In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, Prague, Czech Republic, 2007, pp. 177-180.
- [16] K. Papineni, S. Roukos, T. Ward & W. J. Zhu, "BLEU: a method for automatic evaluation of machine translation," In *Proceedings of the 40th annual meeting on association for computational linguistics*, Philadelphia, Pennsylvania, 2002, pp. 311-318.
- [17] H. Hoang, & P. Koehn, "Design of the mooses decoder for statistical machine translation," In *Software Engineering, Testing, and Quality Assurance for Natural Language Processing*, Columbus, Ohio, 2008, pp. 58-65.
- [18] M. J. Hunt, "Figures of merit for assessing connected-word recognisers," *Speech Communication*, Vol 9, No. 4, pp. 329-336, 1990.
- [19] A. Stolcke, "SRILM-an extensible language modeling toolkit," In *Proceedings of the 7th International international conference on spoken language processing*, Denver, Colorado, 2002, pp. 901-904.
- [20] P. Koehn, "Pharaoh: a beam search decoder for phrase-based statistical machine translation models," In *6th Conference of the Association for Machine Translation in the Americas*, AMTA, 2004, pp. 115-124.
- [21] F. J. Och & H. Ney, "a systematic comparison of various statistical alignment models," *Computational linguistics*, Vol. 29, No. 1, pp. 19-51, 2003.
- [22] F. J. Och, "an efficient method for determining bilingual word classes," In *Proceedings of the ninth conference on European chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, Bergen, Norway, 1999, pp. 71-76.
- [23] G. Doddington, "Automatic evaluation of machine translation quality using n-gram co-occurrence statistics," In *Proceedings of the second international conference on Human Language Technology Research*, Morgan Kaufmann Publishers Inc, 2002, pp. 138-145.
- [24] P. Koehn, "MOSES, Statistical Machine Translation System," *User Manual and Code Guide*, 2010, available at: [www.statmt.org/moses/manual/manual.pdf](http://www.statmt.org/moses/manual/manual.pdf).
- [25] P. Koehn, & J. Schroeder, "Experiments in domain adaptation for statistical machine translation," In *Proceedings of the Second Workshop on Statistical Machine Translation*. Association for Computational Linguistics, Prague, Czech Republic, 2007, pp. 224-227.
- [26] Aaron B. Phillips and Ralf D. Brown. "Training machine translation with a second-order taylor approximation of weighted translation instances", In *MT Summit*, 2011.
- [27] Phillips AB, "Cunei: open-source machine translation with relevance-based models of each translation instance", in *special issue: free/open-source machine translation machine*

translation. Mach Trans Vol. 25, No. 2, pp. 161–177, 2011.