Statistical Machine Translation between Myanmar Sign Language and Myanmar SignWriting

Hnin Wai Wai Hlaing
University of Technology (Yatanarpon Cyber City), UTYCC,
Pyin Oo Lwin, Myanmar
hninwaiwaihlaing.ycc@gmail.com

Ye Kyaw Thu
Okayama Prefectural University, OPU,
Okayama, Japan
Waseda University,
Tokyo, Japan
wasedakuma@gmail.com

Swe Zin Moe
University of Technology (Yatanarpon Cyber City), UTYCC,
Pyin Oo Lwin, Myanmar
swezinmoe.1011@gmail.com

Hlaing Myat Nwe
University of Technology (Yatanarpon Cyber City), UTYCC,
Pyin Oo Lwin, Myanmar
hlaingmyatnwe.nwe@gmail.com

Ni Htwe Aung
Yangon Technological University, YTU,
Yangon, Myanmar
nhadec@gmail.com

Nandar Win Min
University of Technology (Yatanarpon Cyber City), UTYCC,
Pyin Oo Lwin, Myanmar
nandarwinmin@gmail.com

Hnin Aye Thant
University of Technology (Yatanarpon Cyber City), UTYCC,
Pyin Oo Lwin, Myanmar
hninayethant@gmail.com

***Abstract*—This paper contributes the first evaluation of automatic machine translation between Myanmar Sign Language (MSL) and Myanmar SignWriting (MSW). The main motivation is to introduce SignWriting to the Myanmar Deaf society with the help of statistical machine translation. In this paper, we use our MSL-MSW corpus for general domain that contains a textual representation of MSL and its equivalent Myanmar SignWriting. The methods studied in this work were phrase-based, hierarchical phrase-based and the operation sequence model. In addition, two different segmentation schemes were studies, these were syllable segmentation and word segmentation for MSL. The performance of the machine translation systems was automatically measured in terms of BLEU and RIBES for all experiments. Our main findings were that operation sequence model gave the highest scores (37.54 BLEU and 0.8280 RIBES) for MSL to MSW translation and hierarchical phrase based machine translation gave the highest scores (52.79 BLEU and 0.8756 RIBES) for MSW to MSL translation. Generally, translation with word segmented MSL achieved better performance than syllable segmentation of MSL. Our 10-fold cross validation results produced promising results even with the limited training data and we expect this can be developed into a useful machine translation system as more data becomes available in the future.**

***Index Terms*—Machine Translation, Hierarchical Phrase-based Machine Translation (HPBSMT), Myanmar Sign Language (MSL), Myanmar SignWriting (MSW), Operation Sequence Model (OSM), Phrase-based Machine Translation (PBSMT)**

I. Introduction

As reported by the 2014 Myanmar national census, about 1.3 percent of the populations in Myanmar are deaf or hard-of-hearing [1]. Myanmar Sign Language (MSL) is used as a primary communication language among Deaf people. They face various difficulties in communicating with other hearing people and feel isolating from their surroundings because there are limited resources of information written in their language. Most of the Myanmar Deaf people are difficult to read or write the standard Myanmar written language because the grammar structure and usage of Myanmar written language and Myanmar sign language are not the same. Additionally, Myanmar language is a tonal and syllable-based. For these reasons, we wish to break down the language barrier and able to get better communication between hearing-impaired and normal-hearing people. Our first motivation is studying statistical machine translation methodologies between textual representation of Myanmar sign language and SignWriting. In other words, this research is developing SignWriting for Myanmar sign language. From this research, we are expacting Deaf people who can understand

Myanmar Sign Language can start learning SignWriting and develop various written resources of current Myanmar sign language such as MSL-MSW dictionary, MSW educational textbooks.

This paper contributes the first evaluation of automatic machine translation between MSL and MSW in both directions. One more contribution is we are developing a parallel corpus of MSL-MSW on the general domain. In this experiment, we used the current version of the MSL-MSW corpus relating to general domain.

The structure of the paper is as follows. In the next section, we present a brief review of machine translation systems for sign languages including SignWritng. Section III presents a sketch of MSL and common phenomena of many other SLs. Section IV introduce SignWriting and Section V presents the MSL-MSW parallel corpus preparation for machine translation experiments. In Section VI, we describe the methodologies used in the machine translation experiments. Section VII presents statistical information of the corpus and the experimental settings. The results together with some discussions are presented in Section VIII. Section IX presents the error analysis on translated outputs. Final Section X concludes our works and indicates promising results for future research.

II. Machine Translation for Sign Language

Machine translation is an application of computers to translate from one natural language to other languages that convey the meaning of the original source language. An automated sign language machine translation is in great demand to make more information and services accessible to persons with hearing and speaking disabilities on a more economical basis when an interpreter is unavailable.

MT systems between spoken and sign languages had a start in the late 90s. Strategies used for developing MT system are also used for developing text to sign language MT system including direct MT, template-based MT, transfer-based MT, interlingua-based MT, rule-based MT, syntax-based MT and statistical-based MT. Details of each strategy can be found in several books as follows: Hutchins and Somers, 1992 [3]; Hutchins, 2000 [2]; Nirenburg and Raskin, 2004 [4]. A number of sign language machine translation systems have been carried out around the world, e.g. TESSA system (Bangham & Cox, 2000) [5], weather reports generate system (Angus & Smith, 1999) [6], South African sign language machine translation system (Zijl & Barker, 2003) [7], experiments in sign language machine translation using examples (Morrissey & Way, 2006) [8], Morpho-syntax base statistical methods for automatic sign language translation (Stein, Bungeroth, & Ney, 2006) [9] and Arabic Text-to-SignWriting Translation (Almasoud & Al-Khalifa, 2012) [10].

III. Sign Language

Sign Language (SL) is the native language of the Deaf community. As spoken language use throat, nose and mouth as articulators, they can express their needs and the formation of concepts by combining hand shapes, orientation and movement of the hands, arms or body and facial expressions. Wherever vocal communication is impossible, sign language can be used to bridge the gap, for example communication between deaf and mute people.

However, SLs are not at all uniform according to the culture and environments. It is not clear how many SLs there are. Each country has its own native sign language and some have more than one. Although SL differs in different regions, it mainly depends on the basic parts of sign. Not only Manual Features (MFs) but also Non-Manual Features (NMFs) are part of a sign. However, sign languges rarely use NMFs alone to communicate. Fig. 1 shows the structure of a sign language.

Manual features are signs formed by one or both hands in different shapes, locations, movements and orientations to express meanings. Non-manual features contain various facial expressions, head tilting, and shoulder raising, mouthing and similar signals that add to hand sign to describe meanings. And then, it grammatically includes questions, negation, relative clause, boundaries between sentences and arguments structure of some verbs [11]. Similar to American Sign Language (ASL) and British Sign Language (BSL), Myanmar Sign Language use non-manual marking for yes/no questions. They are shown through raised eyebrows and a forward head tilt [12], [13], [14]. Fig. 2 shows an example of MSL sentence "အပူချိန် ဒီဂရီ ဘယ်လောက်လဲ ။" + "NMFs –chin up and raised eyebrows for wh-question". The meaning of the MSL sentence is "အပူချိန် ဘယ်လောက်လဲ ။" in Myanmar language and "What is the temperature?" in English respectively.

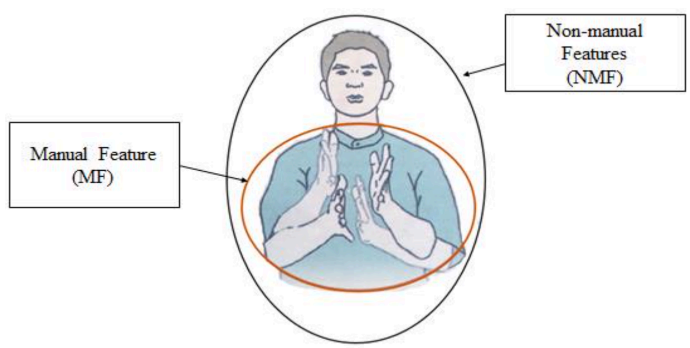


Fig. 1: The structure of sign language

IV. SignWriting

There are many writing systems to represent sign languages in written form in other countries such as Gloss Notation, Hamburg Notation System (HamNoSys) and SignWriting. Among them, SignWriting is becoming widespread because it is language independent, which contain a large number of basic symbols. SignWriting was proposed in 1974 by Valerie Sutton. At first, she is trying to note all dance, all mime and gesture. It is also called movement-writing system for writing all dance, all sports, and all movement. Nowadays, it is becoming SignWriting for Deaf communities. It uses a combination of iconic symbols and the shapes of characters, that are abstract pictures of the hand, body, face and so on [15].

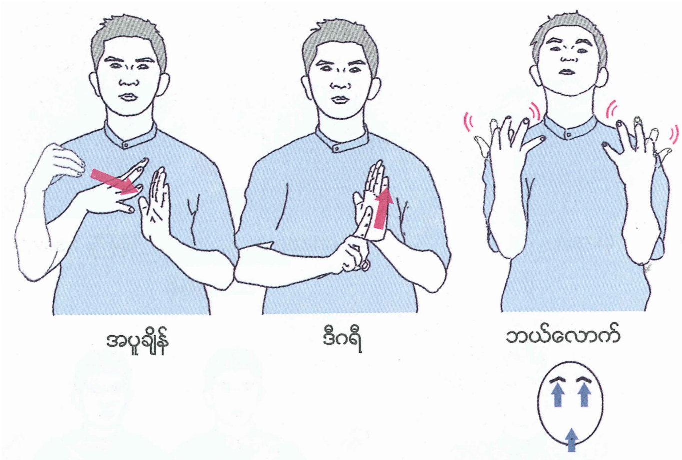


Fig. 2: An example of MSL sentence that used non-manual features (from Myanmar Sign Language Basic Conversation Book)

International SignWriting Alphabet (ISWA) 2010 defines 7 categories, 30 groups of symbols to form 652 base symbols and 35,023 symbols. Fig 3 shows seven categories of ISWA.

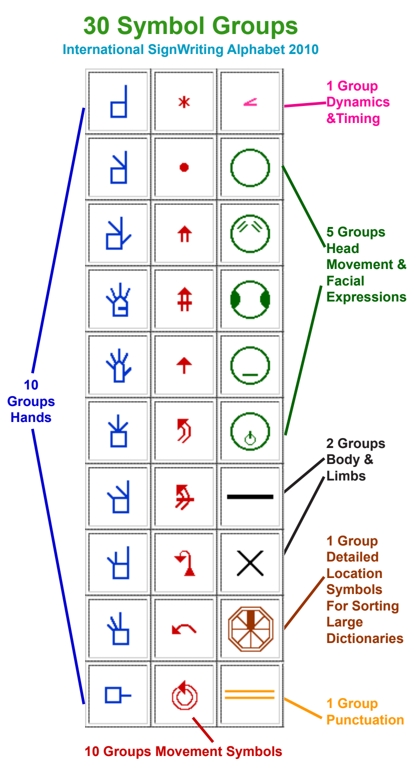


Fig. 3: Seven categories of International SignWriting Alphabet (ISWA), (from the book of "SignWriting Hand Symbols in the International SignWriting Alphabet 2010")

Seven categories of ISWA include hand, movement, dynamic and timing, head and face, body, detailed location, and punctuation. It also represents the sign symbols from both signer's point of view and observer's point of view.

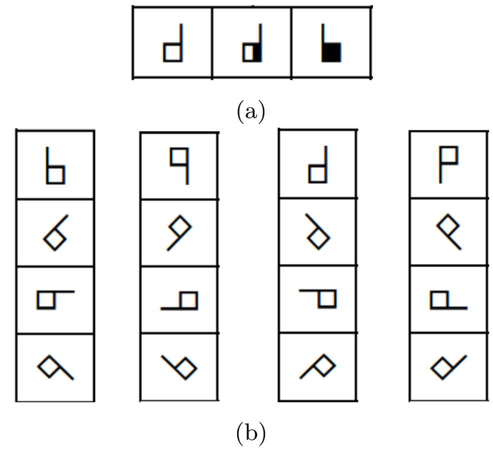


Fig. 4: (a) three different filling symbols (b) sixteen different spatial rotation symbols

However, almost all publications used the signer's point of view and denote the right hand is dominant. These symbols can be arranged both horizontally (left-to-right) and vertically (top-to-bottom), and can be rotated in 8 directions. The orientation of the palm is indicated by filling the glyphs for the hand shape. A white glyph indicates that one is facing the palm of the hand; a black glyph indicates that one is facing the back of the hand and half-shading indicates that one is seeing the hands from the side [16]. Hand orientation is important for SignWriting and there are 3 different filling symbols (see Fig. 4a) and 16 different spatial rotations symbols (see Fig. 4b). SignWriting is the first writing system for sign languages to be included in the Unicode Standard (U+1D800 to U+1DAAF) [17]. The Unicode block for SignWriting is U+1D84C (signwriting hand-flat five fingers spread)-U+1DA9C (signwriting fill modifier-3), which represents five fingers spread seeing the back of the hand from the signer's point of view.

V. Corpus Preparation

Myanmar NLP researchers are facing with many difficulties arising from the lack of resources; in particular parallel corpora are scare [18]. In previous work, we collected 558 Myanmar sign language videos of emergency domain. We also annotated the collected sign language videos with transcription of Myanmar text and translated Myanmar written text [38]. All text sentences are word segmented and POS tagging is done for Myanmar written text [39]. As an extension, in this work, we are building a multimedia parallel MSL-MSW corpus with the purpose of developing a Machine Translation (MT)-based approach for using technology to assist hearing and speaking disabilities with limited Myanmar language in their daily life basic conversation.

In this section, we will explain all the development steps in the construction of MSL-MSW corpus for general domain.

A. Data Collection

The spoken and the written style sentences are manually selected from several Myanmar language books including "the Myanmar Sign Language Basic Conversation Book" [14]. The domain is a general domain, the corpus developing is work in progress and we used 1,448 sentences in this experiment. It contains two main categories and they are MSL for the emergency case and the basic conversation. In details, emergency for fire, earthquake, floods, storms, accident, police and health, and basic conversations for greeting, travelling, occupations, number, date and time. The raw Myanmar sentences to MSL translation was done by discussions with Sign language trainers, native signers and deaf persons to ensure the meaning of the original Myanmar written sentences. After making several discussions, sign language video data are collected for each Myanmar sentence.

This corpus was developed with 22 sign language trainers and Deaf people: males and females, age range from 11 to 48, from School for the Deaf (Mandalay), Mary Chapman School for Deaf Children (Yangon), School for the Deaf (Tarmwe), Myanmar Deaf Society and Literacy and Language Development for the Deaf in Yangon and Mandalay regions has been carried out. The MSL-MSW corpus contains MSL videos, a textual representation or direct annotation of Myanmar sign language with Myanmar language and parallel SignWriting symbols.

B. Manual Annotation with SignWriting

After video data collection had been finished, we have to define SignWriting symbols for each sign of MSL. In details, we watched the recorded video several times for defining both manual and non-manual signs. After that, sign symbols are placed on the canvus of SignMaker to form the shape and movement of signs. SignWriting symbols are needed to arrange in a unique sequence. Fig. 5 shows the sign symbols arrangement in SignMaker. There are two ways to prepare SignWriting data: one is the formal SignWriting–based on 2-dimensional mathematics and written as a string of ASCII characters and another is Unicode representation of SignWriting symbols. In our work, we use Unicode numbers for SignWriting symbols seeing SignWriting symbols arrangement in SignMaker. An example of SignWriting Unicode character sequences for the MSW word "doctor" is as follows and equivalent MSW can be seen in Fig. 5.

English: Doctor
Myanmar: ဆရာဝန်
Unicode Block: "\U1D800\U1DAAA\U1D800\U1DA9C
\U1D80A\U1DA9B\U1DAA8\U1D9FF\U1DA30
\U1D80A\U1DA9B"

C. Syllable and Word Segmentation

In SMT, word segmentation is a necessary step in order to yield a set of tokens upon which the alignment and indeed the whole machine learning process can operate. We did sign unit based word segmentation for both text representation of MSL and MSW. Sign unit based word segmentation was done manually for the whole parallel corpus. It is based on meaningful MSL words considering repeated signs (e.g. two or more repeated "thank you" sign for "please"), sign with multiple meanings (e.g. one MSL sign for "blood" and "red"), compound sign (e.g. combination of MSL signs "car", "emergency" and "fire extinguishing" for "fire truck", name sign (e.g. Yangon city), fingerspelling sign, fingerspelling shortcut sign (Myanmar consonant "မ" (Ma) for Mandalay city) and phrase level signs. Based on the previous studies relating to effectiveness of Myanmar word segmentation schemes for SMT [36], we also decided to use syllable segmentation for MSL. We used Regular Expression (RE) based Myanmar syllable segmentation tool named "sylbreak" [37].

VI. Experimental Methodology

A. Phrase-based Statistical Machine Translation (PBSMT)

A PBSMT translation model is based on phrasal units [19], [20]. Here, a phrase is simply a contiguous sequence of words and generally, not a linguistically motivated phrase. A phrase-based translation model typically gives better translation performance than word-based models. We can describe a simple phrase-based translation model consisting of phrase-pair probabilities extracted from corpus and a basic reordering model, and an algorithm to extract the phrases to build a phrase-table [21].

B. Hierarchical Phrase-based Statistical Machine Translation (HPBSMT)

The hierarchical phrase-based SMT approach is a model [22] based on synchronous context-free grammar. The model is able to be learned from a corpus of unannotated parallel text. The advantage this technique offers over the phrase-based approach is that the hierarchical structure is able to represent the word reordering process. The reordering is presented explicitly rather than encoded into a lexicalized reordering model (commonly used in purely phrase-based approaches). This makes the approach particularly applicable to language pairs that require long-distance reordering during the translation process [23].

C. Operation Sequence Model (OSM)

The Operation Sequence Model (OSM) [24], combines the benefits of phrase-based and N-gram-based SMT [25] and remedies their drawbacks. It is based on minimal translation units, capture source and target context across phrasal boundaries and simultaneously generate source and target units. Providing a strong coupling of lexical generation and reordering gives a better reordering mechanism than PBSMT. The list of operations can be divided into two groups and they are five translation operations (Generate (X, Y), Continue Source Cept, Generate Identical, Generate Source Only (X) and Generate Target Only (Y)) and three reordering operations (Insert Gap, Jump Back (N) and JumpForward).

VII. EXPERIMENTS

A. Corpus Statistics

We used 1,448 Myanmar sign language and Myanmar SignWriting parallel sentences, which is a collection of general domain (refer Section V). In this experiment,

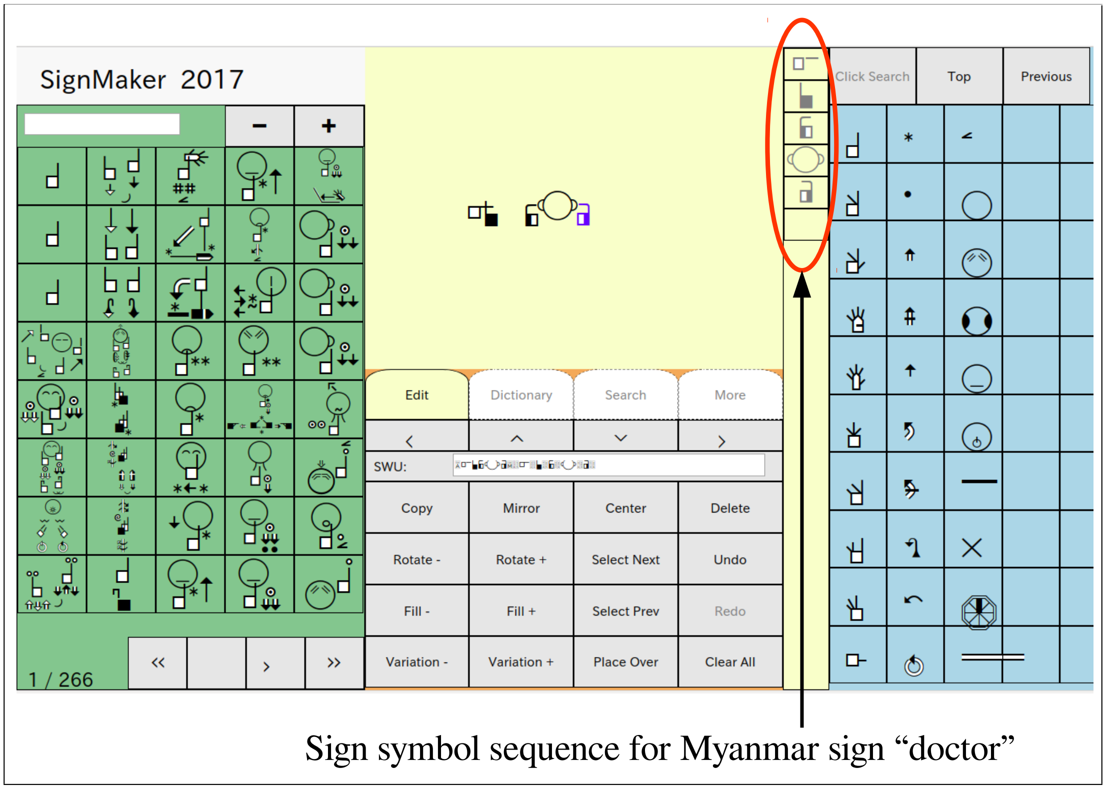


Fig. 5: An example of sign symbol sequence arrangement in the SignMaker 2017

1,000 sentences were used for training, 170 sentences for development and 278 sentences for evaluation.

B. Moses SMT System

We used the PBSMT, HPBSMT and OSM provided by the Moses toolkit [26] for training the PBSMT, HPBSMT and OSM statistical machine translation systems. The word segmented source language was aligned with the word segmented target languages using GIZA++ [27]. The alignment was symmetrized by grow-diag-final-and heuristic [28]. The lexicalized recording model was trained with the msd-bidirectional-fe option [29]. We use KenLM for training the 5-gram language model with interpolated modified Kneser-Ney discounting [30], [31]. Minimum error rate training (MERT) [32] was used to tune the decoder parameters and the decoding was done using the Moses decoder (version 2.1.1) [26]. We used default settings of Moses for all experiments. Our current parallel corpus size is limited and thus 10-fold cross validation was done for all experiments.

C. Evaluation

We used two automatic criteria for the evaluation of the machine translation output. One was the de facto standard automatic evaluation metric Bilingual Evaluation Understudy (BLEU) [33] and the other was the Rank-based Intuitive Bilingual Evaluation Measure (RIBES) [34]. The BLEU score measures the adequacy of the translations and RIBES is suitable for distance language pairs such as Myanmar and English. The higher BLEU and RIBES scores are better.

VIII. Result and Discussion

The BLEU and RIBES score results for machine translation experiments with PBSMT, HPBSMT and OSM between MSL (word) and MSW (word) are shown in Table I. The results for MSL (syllable) and MSW (word) pair are shown in Table II. RIBES scores are shown in brackets. Bold numbers indicate the highest scores among the three SMT approaches.

TABLE I: BLEU and RIBES scores of sign language (word) and SignWriting pair for PBSMT, HPBSMT and OSM

src-trg	Word Segmented MSL		
	PBSMT	***HPBSMT***	***OSM***
MSL-MSW	34.44 (0.8014)	34.99 (0.8049)	**37.54 (0.8280)**
MSW-MSL	52.66 (0.8754)	**52.79 (0.8756)**	49.99 (0.8675)

*the values inside the parentheses are RIBES scores

Looking at the results in Table I and II, MSL(word)-MSW segmentation scheme of MSL was by far the most effective for both MSL-MSW and MSW-MSL translations.

TABLE II: BLEU and RIBES scores of sign language (syllable) and SignWriting pair for PBSMT, HPBSMT and OSM

src-trg	Syllable Segmented MSL		
	PBSMT	***HPBSMT***	***OSM***
MSL-MSW	33.99 (0.8206)	34.04 **(0.8260)**	**34.38** (0.8200)
MSW-MSL	49.47 (0.8660)	49.62 (0.8650)	**50.42 (0.8676)**

*the values inside the parentheses are RIBES scores

In Table. I, MSL-MSW translation achieved the highest BLEU and RIBES scores (37.54 and 0.8280) using OSM approach and MSW-MSL translation gave the highest BLEU and RIBES scores (52.79 and 0.8756) in HPBSMT. From the overall results, it can be clearly seen that OSM approach is better for both MSL to MSW and MSW to MSL translations. PBSMT and HPBSMT results are comparable for both word and syllable segmentations. If we only focus on syllable segmentation experiments (see Table. II), all three SMT approaches, PBSMT, HPBSMT and OSM results are comparable.

IX. Error Analysis

We analyzed the translated outputs using Word Error Rate (WER) [35]. We used the SCLITE (score speech recognition system output) program from the NIST scoring toolkit SCTK version 2.4.10 for making dynamic programming based alignments between reference (ref) and hypothesis (hyp) and calculation of WER. The formula for WER can be stated as equation (1):

$$WER = \frac{(N_i + N_d + N_s) \times 100}{N_d + N_s + N_c} \tag{1}$$

where N_i is the number of insertions, N_d is the number of deletions, N_s is the number of substitutions, N_c is the number of correct words. Note that if the number of insertions is very high, the WER can be greater than 100%. The following example shows WER calculation on the translated outputs of three SMT approaches for MSL-MSW language pair with the word segmentation method. In this example, S=1, D=1, I=1, C=4, N=6 for PBSMT and OSM, its WER is equal to 50%, S=2, D=1, I=1 for HPBSMT and its WER is equal to 66.67 %.

The followings are the annotation of above MSW sentences into MSL for the readers (the underlined words are pointing the differences with the reference):

Ref: မီးသတ် ကား အရေးပေါ် အခု လာမယ် ။
PBSMT Hyp: ကား အရေးပေါ် အခု <u>လွှတ်</u> <u>ရ</u> ။
HPBSMT Hyp: ကား အရေးပေါ် <u>လွှတ်</u> <u>ရ</u> <u>အခု</u> ။
OSM Hyp: ကား အရေးပေါ် အခု <u>လွှတ်</u> <u>ရ</u> ။

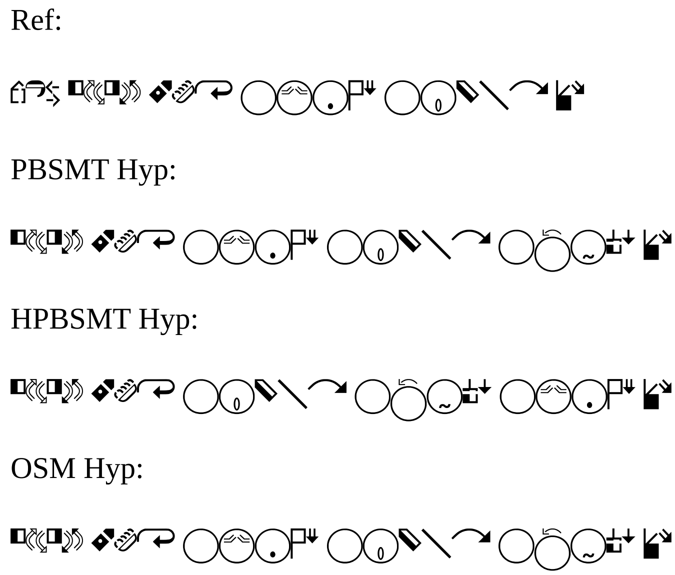


Fig. 6: An example of reference and hypothesis of PBSMT, HPBSMT and OSM

Fig. 7 and Fig. 8 present the WER percentages of translation between MSL and MSW. The results show that "word segmented" MSL gave the lower WER values for both MSL to MSW and MSW to MSL translations.

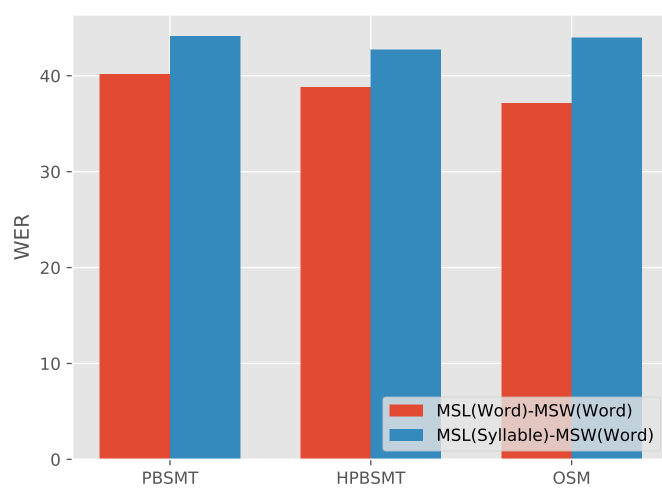


Fig. 7: WER of machine translation from Myanmar sign language to Myanmar SignWriting

From our detail analysis on confusion pairs of three SMT approaches, most of the confusion pairs are caused by the three main reasons and they are (1) the nature of the sign language (2) some errors in the reference or human mistakes (3) limited size of the training data. For example, the top 10 confusion pairs of PBSMT translation model is shown in Table. III. In this table, the 1st column is the reference, the 2nd column is the hypothesis (i.e. output of the PBSMT translation model) and the 3rd column is the description of reference and hypothesis in Myanmar

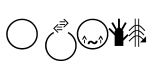
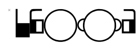
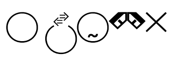
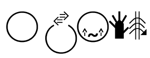
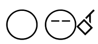
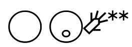
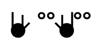
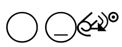
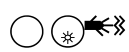
No.	Reference	Hypothesis	Ref-Hyp Description in Myanmar Language	Freq
1			မလုပ်နဲ့ -> မလုပ်နဲ့	7
2			ရထား -> ရထား	4
3			ဆရာဝန် -> ဆရာဝန်	4
4			မသုံးနဲ့ -> မလုပ်နဲ့	3
5			မလုပ်နဲ့ -> မလုပ်နဲ့	3
6			နည်းနည်း -> ခဏ	3
7			လျှပ်စစ် -> လျှပ်စစ်	3
8			ခံလိုက်ရ -> ပြီ	3
9			အလုပ် -> အလုပ်	3
10			ရေ -> ရေ	3

TABLE III: Top 10 confusion pairs of PBSMT model

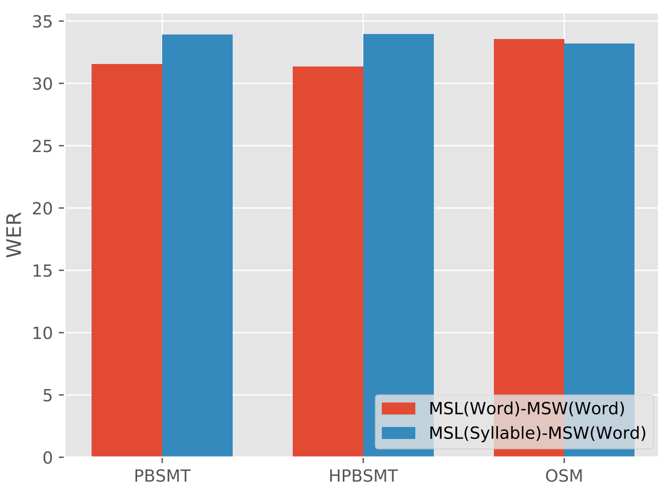


Fig. 8: WER of machine translation from Myanmar SignWriting to Myanmar sign language

written text. Here, confusion pair number 1, 4 and 5 are caused by the same sign language usage for the several meanings. In Myanmar sign language, the word မဟုတ် ("No" in English), the phrase မလုပ်နဲ့ ("Don't do it!" in English), the word မဖြစ် ("do not" in English) and the phrase မသုံးနဲ့ ("Don't use it!" in English) are the same. And thus, the translation model couldn't learn well and we assumed this is also relating to the limited size of our training data. The confusion pair number 2 is caused by the error of the reference data. The confusion pair number 3 and 10 are caused by the sign language dialects (i.e. the difference between Yangon and Mandalay cities sign langauges). One more good example of the confusion pair caused by the sign language nature is the number 7 confusion pair (see Table. III). Although one hand sign is using in the reference of the confusion pair number 7, the hypothesis is using two hands.

X. Conclusion

This paper has presented the first study of the statistical machine translation between Myanmar sign language and Myanmar SignWriting. We implemented three SMT systems (PBSMT, HPBSMT and OSM) with our developing MSL-MSW parallel corpus. We also investigated the effectiveness of two word segmentation schemes (word segmentation and syllable segmentation for Myanmar sign language) for SMT. Our results clearly show that OSM approach achieved the highest translation performance for MSL to MSW translation. However, the HPBSMT approach achieved the highest translation performance for MSW to MSL translation. From our investigation on the effectiveness of word segmentations for MSL-MSW machine translations, the results proved that word segmentation is better than syllable segmentation for MSL side.

In the future work, we plan to expand the MSL-MSW parallel data and conduct experiments on SMT with SignWriting character level (i.e combination of basic symbol, filling symbol and spatial rotation symbol as a one SignWriting character) segmentation approach. We are also doing experiments on Myanmar written text to MSW machine translations.

Acknowledgment

We would like to thank principals, teachers, SL trainers and students of School for the Deaf (Mandalay), Mary Chapman Chapman School for the Deaf Children (Yangon) and School for the Deaf, Tarmwe (Yangon), Myanmar Deaf Society and Literacy and Language Development for the Deaf for their kind contribution to our research. This research is partially supported by Ministry of Education, Department of Higher Education (Myanmar).

References

[1] The population and housing census of Myanmar, 2014
[2] Hutchins, W. J., "Early years in machine translation", John Benjamins Publishing, 2000, USA doi: 10.1075/sihols.97
[3] Hutchins, W. J., and Somers, H. L., "An introduction to machine translation", Academic Press, 1992, London, ISBN-13: 978-0123628305
[4] Nirenburg, S., and Raskin, V., "Ontological semantics", The MIT Press, 2004 ISBN: 9780262140867
[5] Bangham, J. A., and Cox, S. J., "Signing for the deaf using virtual humans", In Proceeding of the Speech and Language Processing for Disabled and Elderly People (Ref. No. 2000/025), IEE Seminar, 2000, London, UK
[6] Angus B., and Smith, G., "English to American sign language machine translation of weather reports", In Proceeding of the 2nd high desert student conference in linguistics (HDSL2), 1999, Albuquerque, New Mexico, pp. 23–30
[7] Zijl, L. V., and Barker, D., "South African sign language machine translation system", In Proceeding of the 2nd international conference on computer graphics, virtual reality, visualisation and interaction in Africa (ACM SIGGRAPH), 2003, Cape Town, South Africa, pp. 49–52
[8] Morrissey, S. and Way, A., "Experiments in sign language machine translation using examples", In Proceeding of the IBM CASCON 2006 Dublin symposium, Dublin, Ireland.
[9] Stein, D., Bungeroth, J., and Ney, H., "Morpho-syntax based statistical methods for sign language translation", In Proceeding of the 11th annual conference of the European association for machine translation, June, 2006, Oslo, Norway.
[10] Almasoud, Al-Khalifa., "Semantic system for Arabic Text-to-SignWriting Translation", Journal of Software Engineering and Applications, 2012,5,604-612
[11] Thompson, RobYin; Emmorey, Karen; Kluender, Robert, "The Relationship between Eye Gaze and Verb Agreement in American Sign Language: An Eye-tracking Study". Natural Language and Linguistic Theory. 24 (2), 2006, pp. 571–604 doi:10.1007/s11049-005-1829-y.
[12] Baker, Charlotte, and Dennis Cokely, "American Sign Language: A teacher's resource text on grammar and culture", 1980, Silver Spring, MD: T.J. Publishers.
[13] Sutton-Spence, Rachel, and Bencie Woll, "The linguistics of British Sign Language", Cambridge: Cambridge University Press, 1998
[14] "Myanmar Sign Language Basic Conversation Book", Ministry of Social Welfare, Relief and Resettlement, Department of Social Welfare, Japan International Cooperation Agency, 1st Edition, August 2009, Daw Yu Yu Swe, Department of Social Welfare
[15] SignWriting Site: http://www.signwriting.org/read.html
[16] Valerie Sutton and Adam Frost., "SignWriting Hand Symbols in ISWA2010: Manual 2"
[17] Sutton SignWriting Unicode., http://www.unicode.org/charts/PDF/U1D800.pdf
[18] Ye Kyaw Thu, Vichet Chea, Andrew Finch, Masao Utiyama and Eiichiro Sumita, "A Large-scale Study of Statistical Machine Translation Methods for Khmer Language", 29th Pacific Asia Conference on Language, Information and Computation, October 30 - November 1, 2015, Shanghai, China, pp. 259-269
[19] Kohen, P., Och, F .J., Marcu, D., "Statistsical phrase-based translation", In HLT-NAACL, 2003, url: http://acl.ldc.upenn.edu/N/N03/N03-1017.pdf
[20] Och, F .j., Marcu, D., Statistical phrase-based translation, 2003, p.127-133.
[21] Specia, L.. "Tutorial, fundamental and new approaches to statistical machine translation", In: International Conference Recent Advances in Natural Language Processing, 2011
[22] Chiang, D.. "Hierarchical phrase-based translation", Comput Linguist 2007;33(2):201-228. url: http://dx.doi.org/10.1162/coli.2007.33.2.201. doi:10.1162/coli.2007.33.2.201.
[23] Braune, F., Gojun, A., Fraser, "A. Long-distance reordering during search for hierarchical phrase-based SMT", In: EAMT 2012: Proceedings of the 16th Annual Conference of the European Association for Machine Translation, Trento, Italy, Citeseer; 2012, pp. 177-184
[24] Durrani, N. Schmid, H., Fraser, A.M., "A joint sequence translation model with integrated reordering". In: Lin, D., Matsumoto, Y., Mihalcea, R., editors. ACL. The Association for Computer Linguistics. ISBN 978-1-932432-87-9;2011, pp. 1045-1054 url: http://dblp.uni-trier.de/db/conf/acl/acl2011.html#DurraniSF11
[25] Mari o, J.B., Banchs, R.E., Crego, J.M., de Gispert, A., Lambert, P., Fonollosa, J.A.R., et al, "N-gram-based machine translation", Comput Linguist 2006;32(4):527-549. url: http://dx.doi.org/10.1162/coli.2006.32.4.527. Doi:10.1162/coli.2006.32.4.527.
[26] Philipp Koehn and Barry Haddow. 2009. "Edinburgh's Submission to all Tracks of the WMT2009 Shared Task with Reordering and Speed Improvements to Moses". In Proceedings of the Fourth Workshop on Statistical Machine Translation, pages 160–164.
[27] Och, F.J., Ney, H.. "Improved statistical alignment model", In ACL00. Hong Kong, China, 2000, pp. 440-447
[28] Koehn, P., Och, F.J.., Marcu, D.., "Statistical phrase-based translation", In Proceedings of the Human Language Technology Conference, 2003, Edmonton, Canada, pp. 48-54
[29] Tillmann, C.., "A unigram orientation model for statistical machine translation", In Proceedings of HLT-NAACL 2004: Short Papers; HLT-NAACL-Short'04. Stroudsburg, PA, USA: Association for Computational Linguistics. ISBN 1-932432-24-8; 2004, pp.101-104, http://d1.acm.org/citation.cfm?id-1613984.1614010.
[30] Heafield, Kenneth, "KenLM: Faster and Smaller Language Model Queries", Proceedings of the Sixth Workshop on Statistical Machine Translation; WMT '11, 2011, Association for Computational Linguistics, Edinburgh, Scotland, pp. 187-197 ISBN- 978-1-937284-12-1
[31] Chen, S.F., Goodman, J., "An empirical study of smoothing techniques for language modeling", In Proceedings of the 34th annual meeting on Association for Computational Linguistics. Association for Computational Linguistics; 1996, pp. 310-318.
[32] Och, F.J., "Minimum error rate training for statistical machine translation", In Proceedings of the 41st Meeting of the Association for Computational Linguistics (ACL 2003). Sapporo, Japan
[33] Papineni, K., Roukos, S., Ward, T., Zhu, W., "Bleu: a Method for Automatic Evaluation of Machine Translation". IBM Research Report rc22176 (w0109022), 2001, Thomas J. Watson Research Center, In ACL '02 Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, July 07 - 12, 2002, Philadelphia, Pennsylvania, pp. 311-318
[34] Isozaki, H., Hirao, T., Duh, K., Sudoh, K., Tsukada, H., "Automatic evaluation of translation quality for distant language pairs", In Proceedings of the Conference on Empirical Methods in Natural Language Processing. pp. 944–952.
[35] Wikipedia of Word Error Rate: https://en.wikipedia.org/wiki/Word_error_rate
[36] Ye Kyaw Thu, Andrew Finch, Yoshinori Sagisaka and Eiichiro Sumita, "A Study of Myanmar Word Segmentation Schemes for Statistical Machine Translation", In Proceedings of the 11th International Conference on Computer Applications (ICCA 2013), February 26-27, 2013, Yangon, Myanmar, pp. 167-179
[37] https://github.com/ye-kyaw-thu/sylbreak
[38] https://github.com/ye-kyaw-thu/MSL4Emergency
[39] Khin War War Htike, Ye Kyaw Thu, Zuping Zhang, Win Pa Pa, Yoshinori Sagisaka and Naoto Iwahashi, "Comparison of Six POS Tagging Methods on 10K Sentences Myanmar Language (Burmese) POS Tagged Corpus", at 18th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2017), April 17 23, 2017, Budapest, Hungary