Statistical Machine Translation between Myanmar Written Text and Myanmar SignWriting

Presented by - **Hnin Wai Wai Hlaing**

M.E(IST)-23

Supervisor - Dr. Nandar Win Min

Co-supervisor - Dr. Ye Kyaw Thu, Daw Swe Zin Moe

Outlines

- Abstract
- Motivation
- Objectives
- Sign Language
- SignWriting
- Corpus Preparation
- Segmentation
- Experiments
- Error Analysis
- Conclusion
- Future Work

Abstract

- The research contributes the first evaluation of the quality of automatic translation between Myanmar Written Text (MWT) and Myanmar SignWriting (MSW).
- Myanmar written text (MWT), Myanmar sign language (MSL) and Myanmar SignWriting (MSW) parallel corpus was used for translation.
- Three translation experiments were done for **MWT-MSW, MWT-MSL and MSL-MSW**, in both directions.
- Experiments were carried out using three different statistical machine translation (SMT) approaches: **Phrase-based**, **Hierarchical phrase-based** and **Operation sequence model**.
- Three different segmentation schemes were applied: **Syllable segmentation**, **Word segmentation** and **Sign unit based segmentation** .
- It can solve the difficulties for Deaf in daily life and help them especially in **emergency situations**.

Motivation

- According to 2014 Myanmar National census, about 1.3 percent of population in Myanmar are Deaf or hard-of-hearing people.
- They face various difficulties in communicating with hearing people and feel isolating from their surroundings.
- There are limited resources of information written in their language (e.g., SignWriting Wiki, SignWriting book, SignWriting dictionary).
- The main motivation is to introduce SignWriting to the Myanmar Deaf society using statistical machine translation.
- This will reduce the gap between the Deaf and hearing people of the country.

Objectives

- To learn Machine Translation between Myanmar Written Text (MWT) and Myanmar SignWriting (MSW)
- To develop MWT, MSL, MSW parallel corpus
- To measure machine translation performance using Statistical Machine Translation (SMT) approaches
- To introduce SignWriting to the Myanmar Deaf society
- To fulfill the communication requirements between Deaf and hearing people

Sign Language (SL)

- SL is the native language of the Deaf community.
- Deaf 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**.
- It is not a universal language –each country has its own, native sign language according to their culture.
- Myanmar sign language (MSL) has its own grammar structure which is very difference with Myanmar written text (MWT).

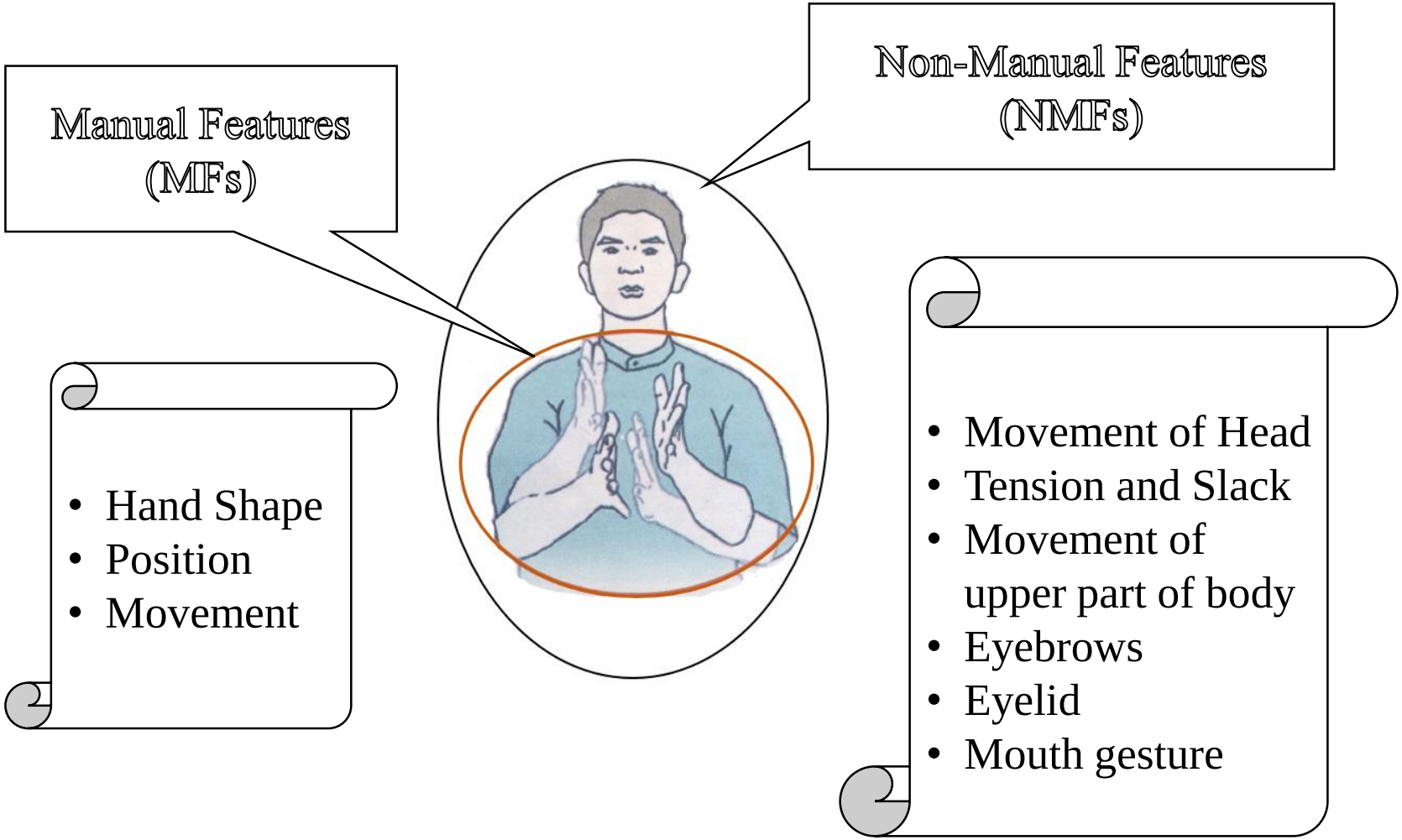


Figure 1 : Structure of Sign Language

SignWriting (SW)

- SW was developed in 1974 by Valerie Sutton, dancer and movement analyst.
- It is a writing system that is a sequence of symbols for sign language.
- Deaf represents two perspective: **signer's perspective** and **observer's perspective**.

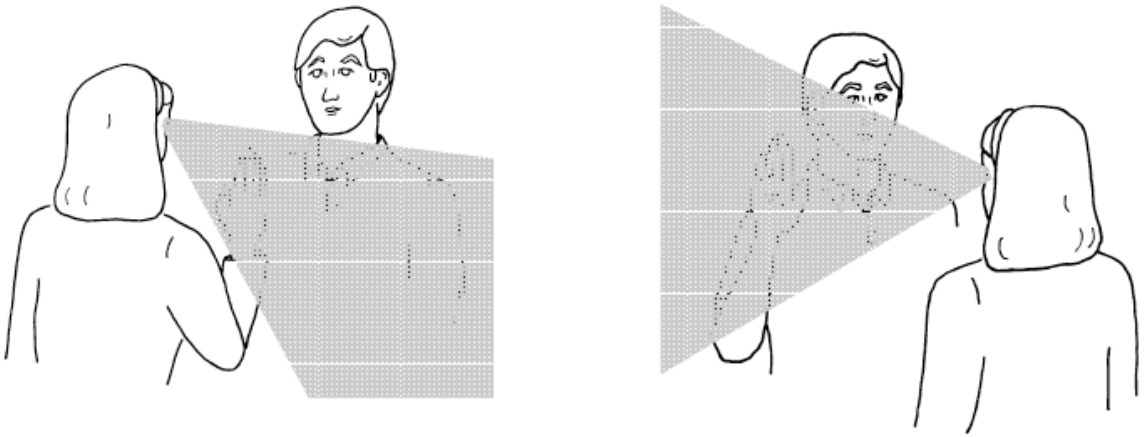
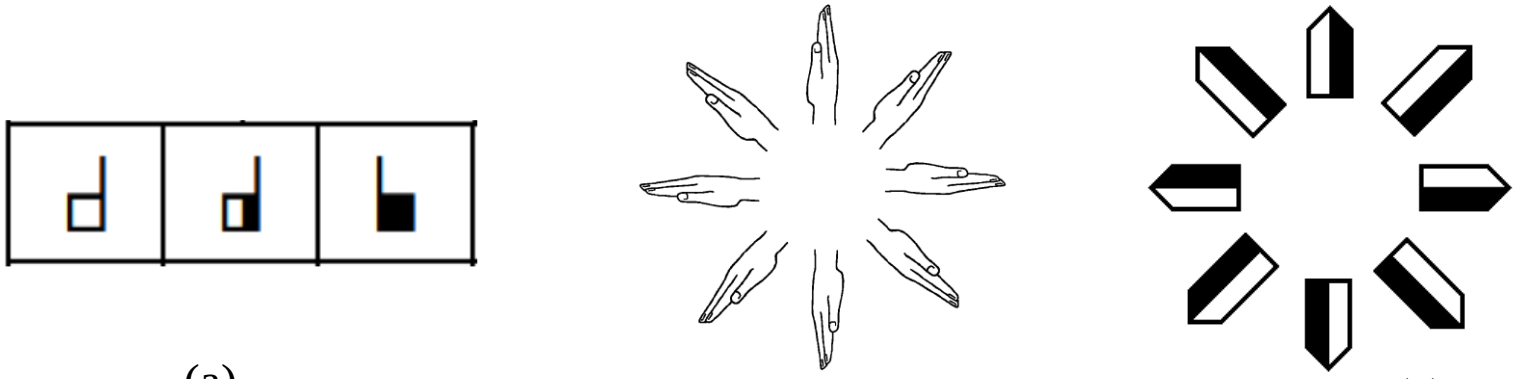


Figure 2: (a) Signer's perspective, (b) Observer's perspective

SignWriting (SW)

- It is written horizontally (left to right) and the right hand is dominant.
- International Sign Writing Alphabet (ISWA) 2010 defines 7 categories, 30 groups of symbols to form 652 base symbols and 35,023 final symbols.
- Hand orientation is important for SignWriting and there are **3 different filling symbols** and **8 different spatial rotations** symbols for each hand.

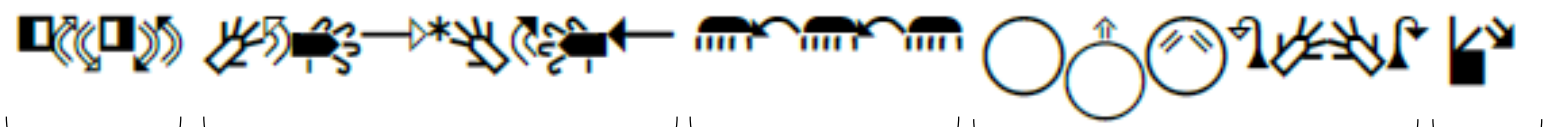


(a) (b) (c)

Figure 3: (a) three filling symbols, (b) eight rotations of hand and (c) its SignWriting symbols

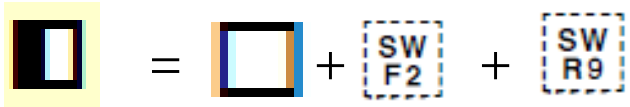
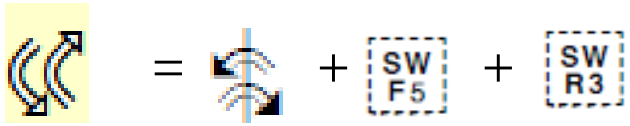
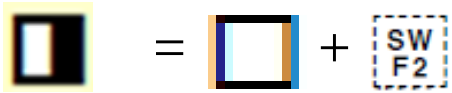
MWT, MSL, MSW parallel sentences

- Example of different grammar, syntax, vocabulary of MWT, MSL and MSW can be seen as follows:
  - Myanmar Written Text (MWT) : ကား တိုက် တဲ့ နေရာ ဘယ် မှာ လဲ ။
  - Myanmar Sign Language (MSL) : ကား ကားတိုက် နေရာ ဘာလဲ ။
  - Myanmar SignWriting (MSW) :



ကား ကားတိုက် နေရာ ဘာလဲ ။

SignWriting (SW)

- An example of SignWriting data preparation is shown as follow:
- Representation of "ന്ദഃ" is 
  - Left fist 
  - Movement of Left fist 
  - Right fist 
  - Movement of Right fist 

System Design

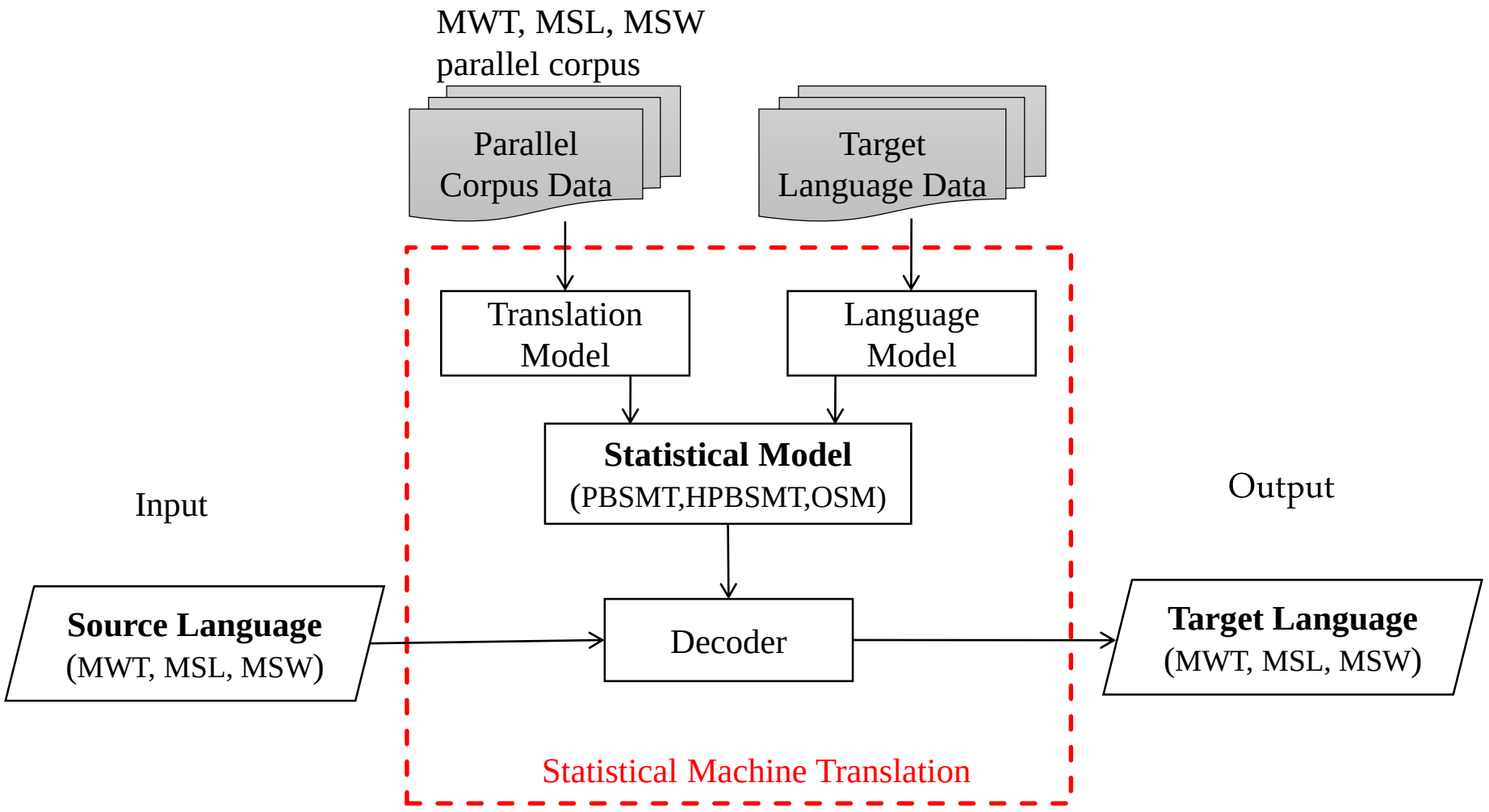


Figure 4: System Design of the proposed system

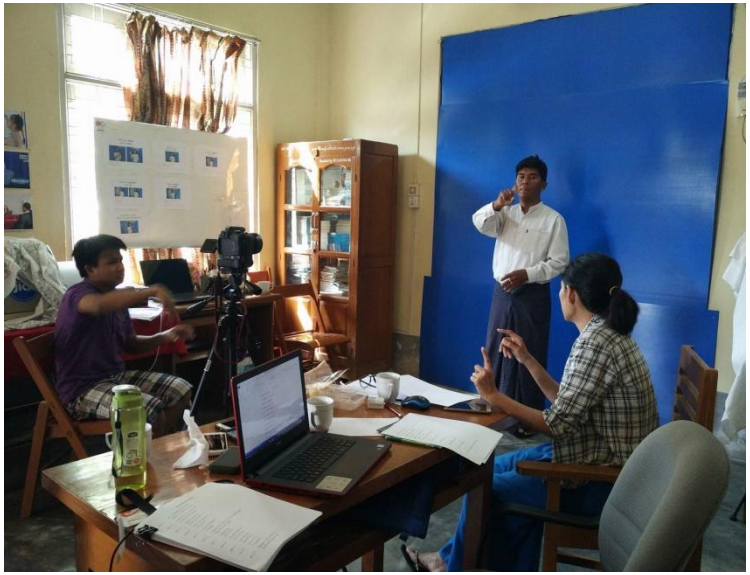
Corpus Preparation

- The corpus contains Myanmar written text (MWT), transcribed Myanmar sign language (MSL) and the Myanmar SignWriting (MSW).
- Three steps are needed for corpus preparation:
  1. Video data collection
  2. Manual Annotation with SignWriting
  3. Segmentation

Video Data Collection

- **Selection of Myanmar Sentence**

Video Data Collection

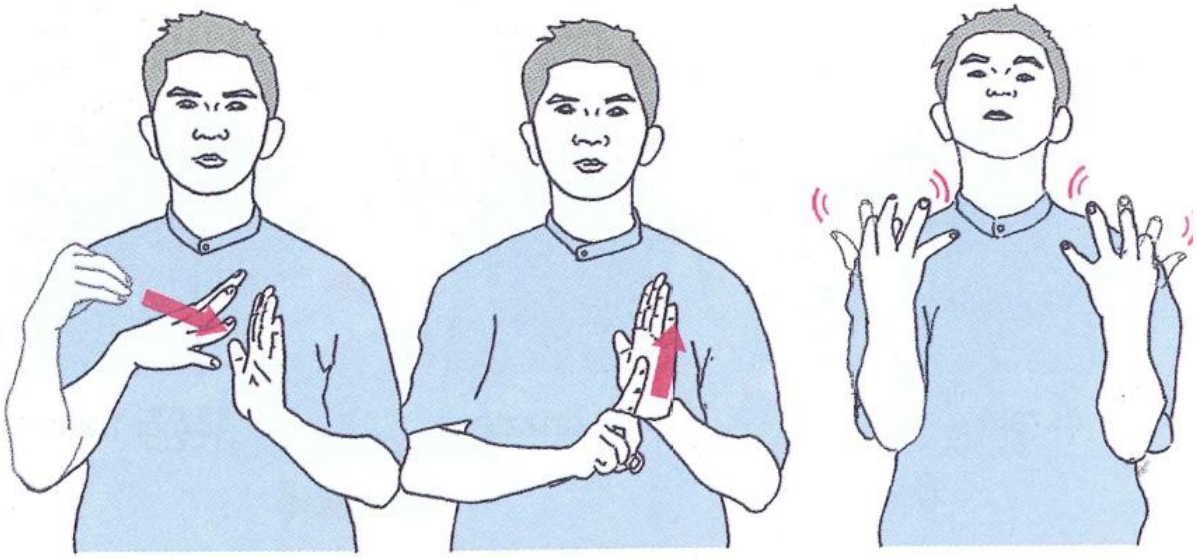


- Selection of Myanmar Sentence
- **Discussion with Sign Language Trainers and Native Signers**

Video Data Collection

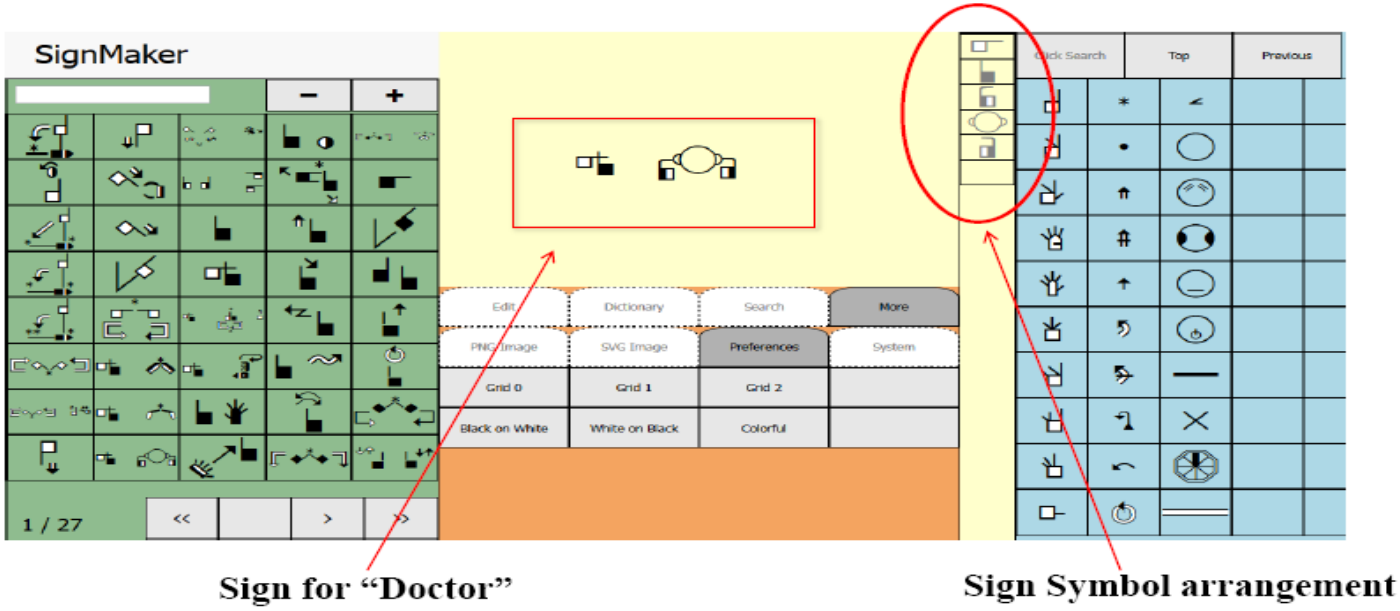
- Selection of Myanmar Sentence
- Discussion with Sign Language Trainers and Native Signers
- **Video Recording**

Manual Annotation with SignWriting



- **Looking each sign in video (Non-manual and Manual Signs)**

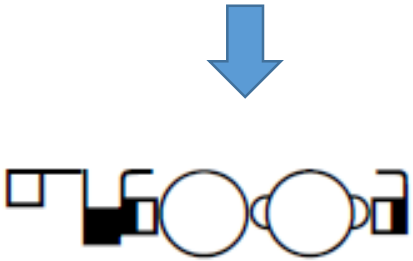
Manual Annotation with SignWriting



- Looking each sign in video (Non-manual and Manual Signs)
- **Sign symbols arrangement in SignMaker**

Manual Annotation with SignWriting

\U1D800\U1DAAA\U1D800\U1DA9C\U1D80A\U1DA9B\U1DAA8\U1D9FF\U1DA30\U1D80A\U1DA9B



- Looking each sign in video (Non-manual and Manual Signs)
- Sign symbols arrangement in SignMaker
- **Defining Unicode sequence and converting SignWriting symbols**

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 learning process can operate.
- Based on the previous studies of Myanmar word segmentation schemes, three segmentation schemes are used for MWT and MSL. They are:
  - Syllable Segmentation
  - Word segmentation
  - Sign Unit based segmentation

Syllable Segmentation

- Basic units for pronunciation of Myanmar words
- Consonant based syllables –describes with Backus Normal Form(BNF) as follows:

Syllable := C{M}{V}[CK][D]

- C for consonants, M for medials, V for vowels, K for vowel killer character and D for diacritic character
- Myanmar syllable segmentation can be done with rule-based, finite state automation(FSA) or regular expression (RE).
- This work use RE based Myanmar syllable segmentation tool, named “sylbreak”

https://github.com/ye-kyaw-thu/sylbreak

Word Segmentation

- In Myanmar text, spaces are used for separating phrases for easier reading.
- There are no clear rules for using spaces in Myanmar text, and thus spaces may (or may not) be inserted between words, phrases.
- In this corpus, manual word segmentation was done for MWT.

Sign Unit based Segmentation

- For MSL sentences, segmentation is based on meaningful MSL word like other sign languages such as American sign language (ASL).
- Some examples of MSL word category are:
  - Repeated sign(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(e.g. “O” sign + “2” sign for “O_2”)
  - Fingerspelling shortcut sign (“O” for Octane, Myanmar consonant “မ” (Ma) for Mandalay city)
  - Phrase or sentence level signs (e.g. MSL sign for စိတ်ငြိမ်ငြိမ်ထား(calm down) , ကားတိုက်(car accident))

Segmentation

- The different segmentation of MWT and MSL

("What is the temperature?" in English)

- **Syllable Segmentation for Myanmar Written Text:**
  အ ပူ ချိန် ဘယ် လောက် လဲ ။

- **Word Segmentation for Myanmar Written Text:**
  အပူချိန် ဘယ်လောက်လဲ ။

- **Sign Unit Based Segmentation for Myanmar Sign Language:**
  အပူချိန် ဒီဂရီ ဘယ်လောက်လဲ ။

Statistical Machine Translation (SMT)

- SMT approach can be described as modeling the probability distribution P(e|f), where e is a string in the source language and f is a string in the target language.
- Using Bayes' Rule, this can be rewritten as;

$$\hat{e} = \text{argmax}_e \, P(e|f) = \text{argmax}_e \, P(f|e) \, P(e)$$

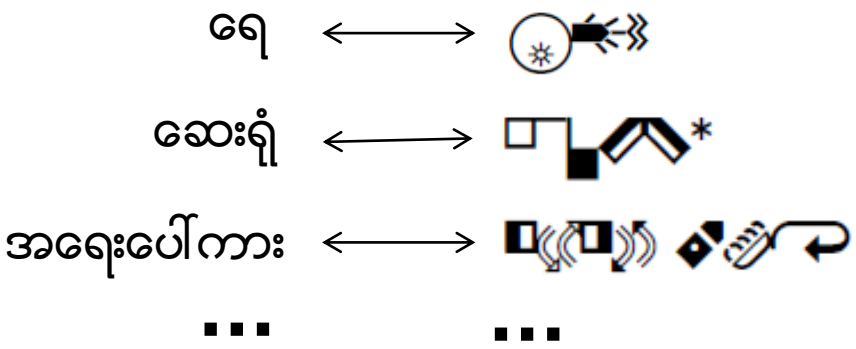
  - P(f|e) : Translation Model
  - P(e) : Language Model
  - argmax_e : Decoder
- Three different statistical machine translation approaches are used: **phrase-based SMT (PBSMT), hierarchical phrase-based SMT (HPBSMT) and operation sequence model (OSM).**

Phrase-Based Statistical Machine Translation (PBSMT)

- PBSMT translates phrases as atomic units and better translation performance than word-based.
- Phrase is a continuous sequence of words and not necessarily a linguistic phrase.
- PBSMT consists of
  - Phrase-pair probabilities extracted from corpus,
  - Reordering model, and
  - An algorithm to extract the phrases to build a phrase-table.

Phrase-Based Statistical Machine Translation (PBSMT)

- First stage in training a phrase-based (PB) model is extraction of PB lexicon
- A PB lexicon pairs strings in one language with string in another language, e.g.,



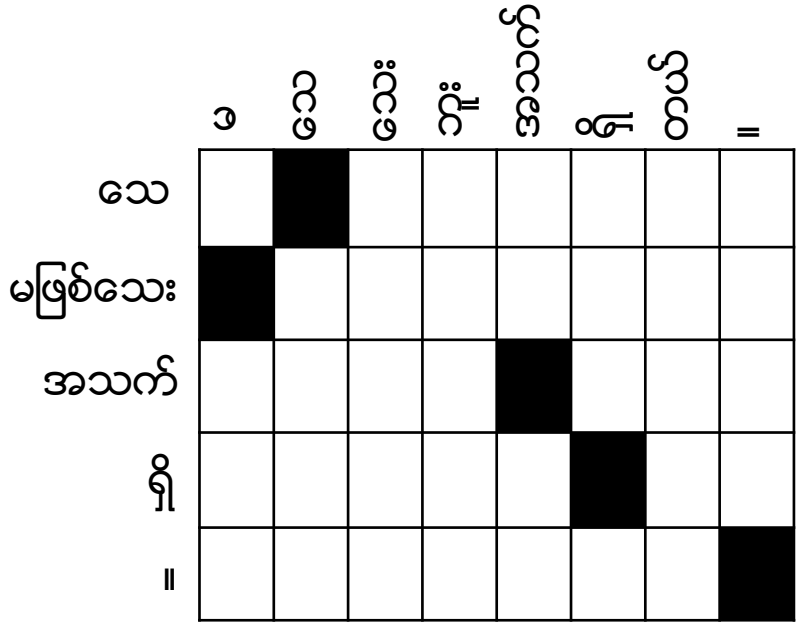
Phrase-Based Statistical Machine Translation (PBSMT)

Finding Alignment Matrices

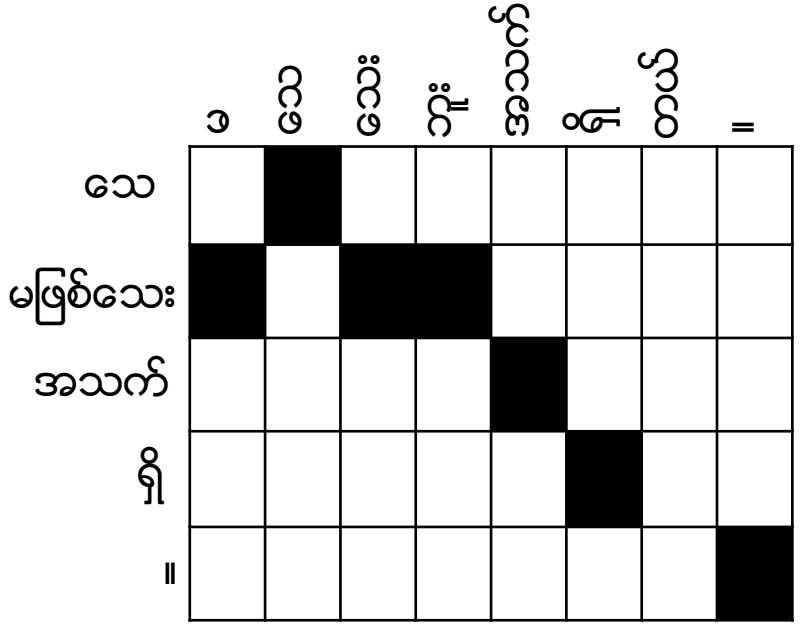
- Step 1: train IBM Model 2 for P($f|e$), and find the most likely alignment for each (e, f) pair
- Step 2: train IBM Model 2 for P($e|f$), and find the most likely alignment for each (f, e) pair
- Given the two alignments, take the intersection of the two as a starting point.

Phrase-Based Statistical Machine Translation (PBSMT)

- **Finding Alignment Matrices**



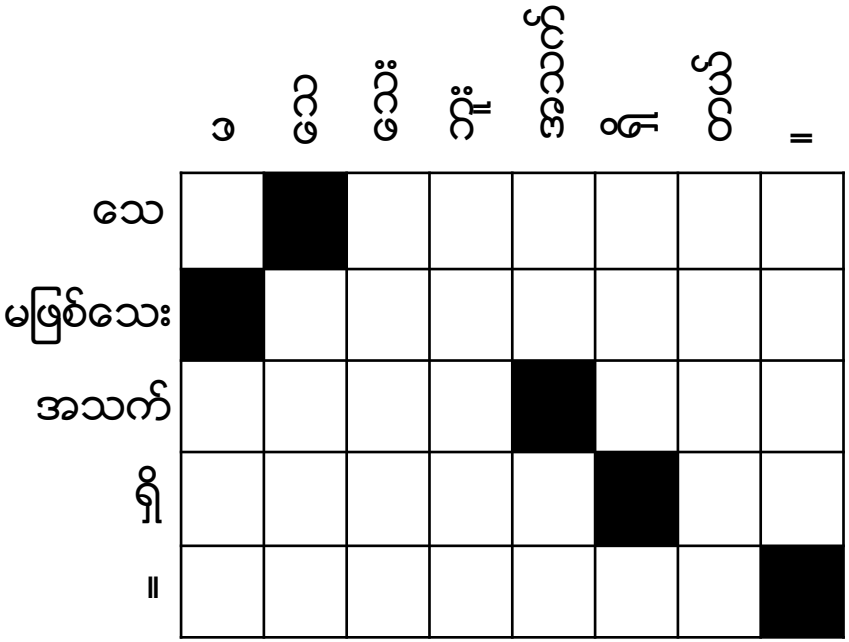
Alignment from P(f | e)



Alignment from P(e | f)

Phrase-Based Statistical Machine Translation (PBSMT)

- **Finding Alignment Matrices**



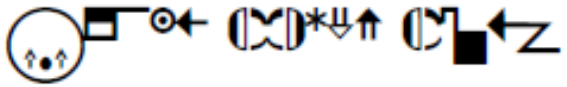
Intersection of Two Alignments

Phrase-Based Statistical Machine Translation (PBSMT)

- **Phrase Translation Probability**
  - For any phrase pair (f,e) extracted from the training data, can calculate

$$Ø(f|e) = \frac{count(f,e)}{count(e)}$$

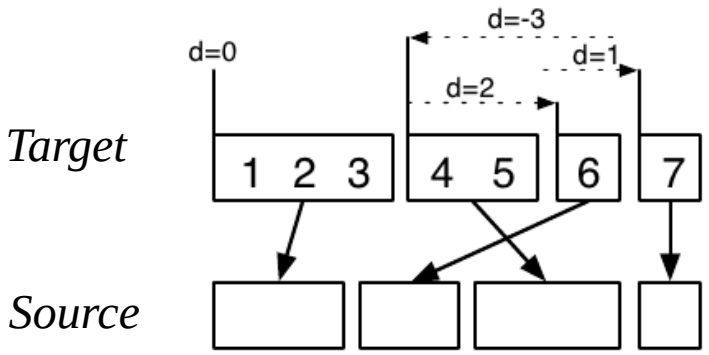
  - For example:

မီးသတ်ဆေးဘူး =  (အနီ ဘူး ဖြန်း)

$$Ø(\text{အနီ ဘူး ဖြန်း} \mid \text{မီးသတ်ဆေးဘူး}) = \frac{count(\text{အနီ ဘူး ဖြန်း , မီးသတ်ဆေးဘူး})}{count(\text{မီးသတ်ဆေးဘူး})}$$

Phrase-Based Statistical Machine Translation (PBSMT)

- **Distance-based Reordering Model**



phrase	translates	movement	distance
1	1–3	start at beginning	0
2	6	skip over 4–5	+2
3	4–5	move back over 4–6	-3
4	7	skip over 6	+1

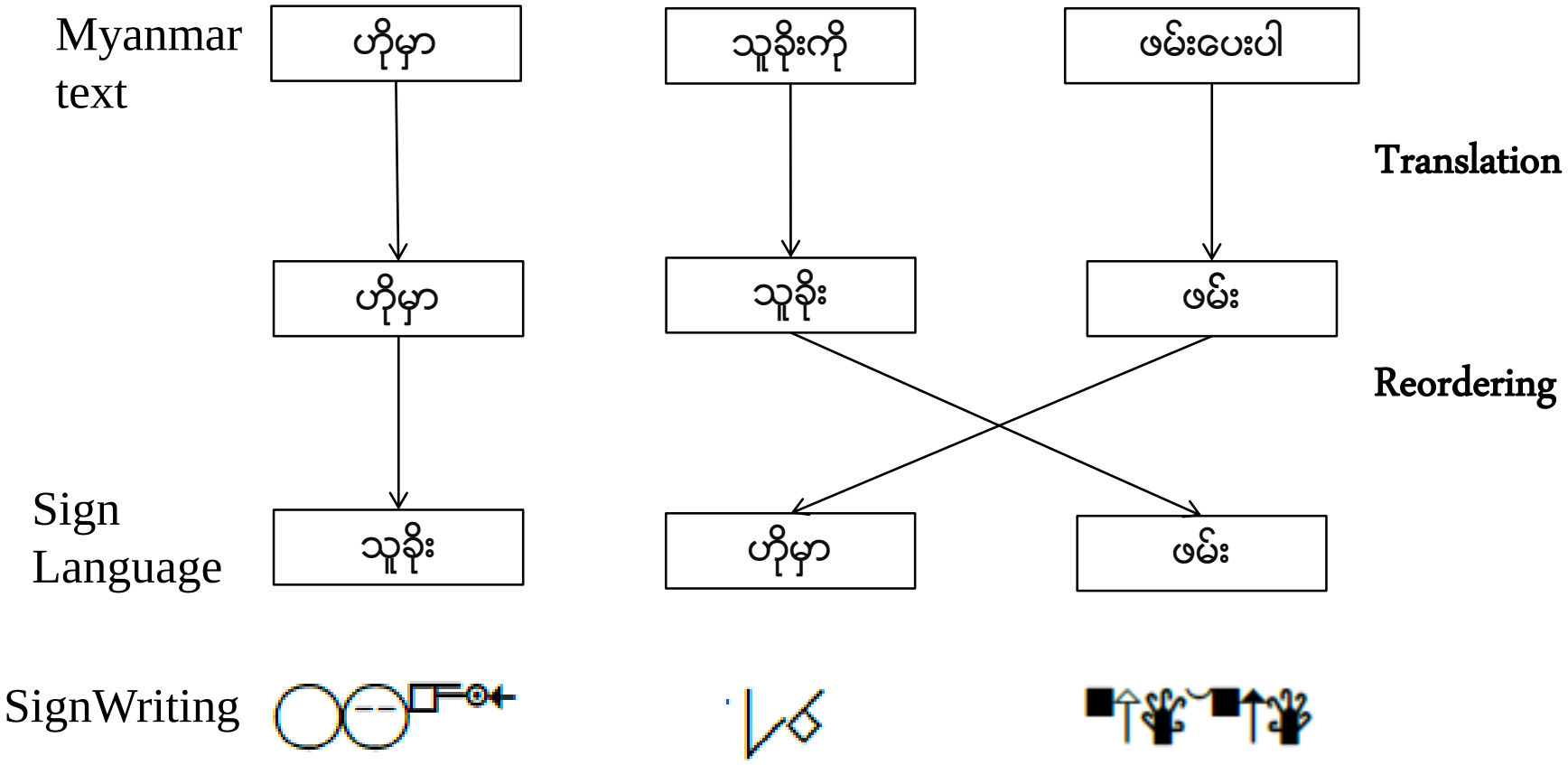
- distance = $start_i$ – end_{i-1} - 1

Phrase-Based Statistical Machine Translation (PBSMT)

- Phrase Translation for "သေ" ("die" in English)

Myanmar	P(e |f)
သေ	0.102086
မ သေ	0.0113429
သေ ရှိ	0.0510431
သေ ရှိ မရှိ	0.0695603
သေ ရှိ မရှိ ဘာလဲ	0.0113429
သေ ရှိ မရှိ ဘာလဲ ။	0.0510431
....	….

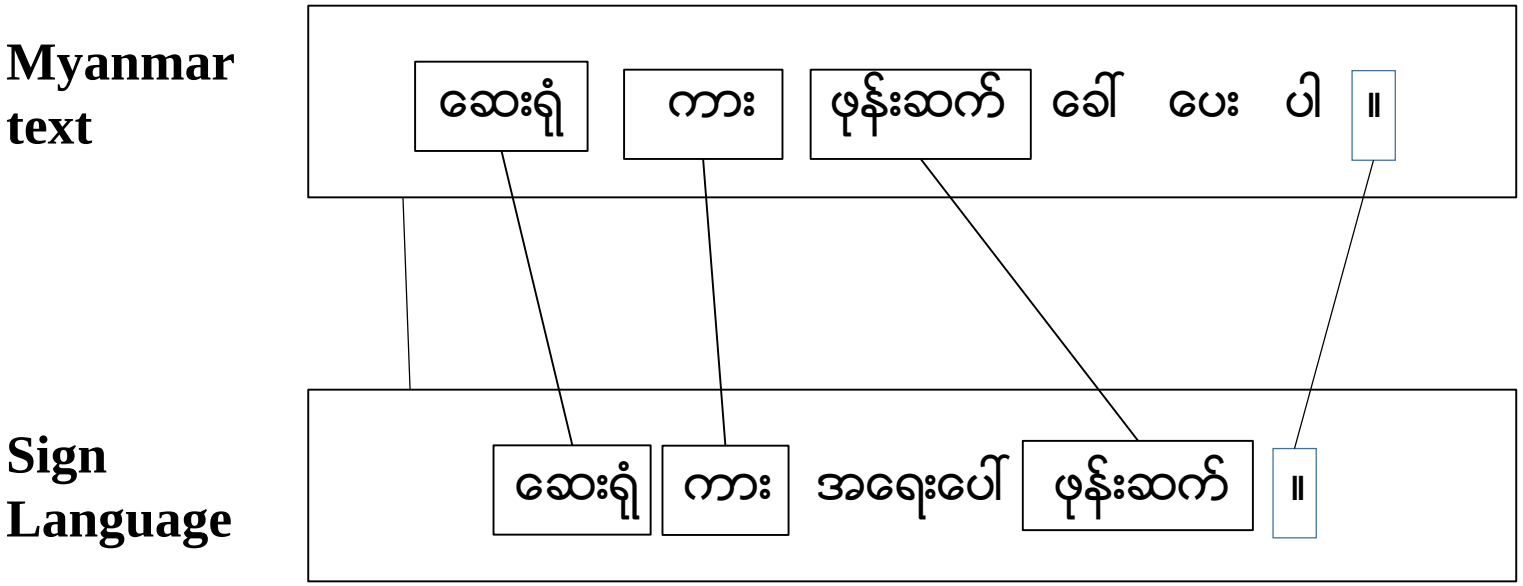
Phrase-Based Statistical Machine Translation (PBSMT)



Hierarchical Phrase-Based Statistical Machine Translation (HPBSMT)

- HPBSMT is a model based on synchronous context-free grammar.
- It learns from a corpus of unannotated parallel text.
- Its advantage over PBSMT is able to represent word reordering process.
- The reordering is represented explicitly rather than encoded into a lexicalized reordering model.
- It is applicable to language pairs that require long-distance reordering during translation process.

Hierarchical Phrase-Based Statistical Machine Translation (HPBSMT)



$$X \longrightarrow \{X1\ X2\ X3\ \text{ခေါ်}\ \text{ပေး}\ \text{ပါ}\ X4 \mid X1\ X2\ \ \text{အရေးပေါ်}\ X3\ X4\}$$

Operation Sequence Model (OSM)

- It combines the benefits of phrase-based and N-gram-based SMT and remedies their drawbacks.
- List of Operations can be divided into two groups.

- **Five Translation Operations**
  - Generate (X,Y)
  - Continue Source Cept
  - Generate Identical
  - Generate Source only (X)
  - Generate Target only (Y)
- **Three Reordering Operation**
  - Insert Gap
  - Jump Back (N)
  - Jump Forward .

Source：ဆေးဆိုင် ကို သွား ပြီး ဆေး ဝယ် လိုက် ပါ ။
Target： ဆေးဆိုင် အဲ့ဒီမှာ ဆေး ဝယ် ။

Operations	Generation	States
		i=0, j=0, k=0 Z=0, j'=0
Generate(ဆေးဆိုင်,ဆေးဆိုင်)	ဆေးဆိုင် ↓ ဆေးဆိုင်	i=1, j=1, k=0 Z=1, j'=1
Generate(ကို, အဲ့ဒီမှာ), Generate Source Only(သွား)	ဆေးဆိုင် ကို ↓ ↓ ဆေးဆိုင် အဲ့ဒီမှာ	i=2, j=3, k=0 Z=3, j'=4
Insert Gap, Generate(ဆေး, ဆေး)	ဆေးဆိုင် ကို ▭ ဆေး ↓ ↓ ↙ ဆေးဆိုင် အဲ့ဒီမှာ ဆေး	i=3, j=5, k=0 Z=5, j'=5
Generate(ဝယ်, ဝယ်), Generate Source Only(လိုက်)	ဆေးဆိုင် ကို ▭ ဆေး ဝယ် ↓ ↓ ↙ ↙ ဆေးဆိုင် အဲ့ဒီမှာ ဆေး ဝယ်	i=4, j=7, k=0 Z=7, j'=8
Insert Gap, Generate(။, ။)	ဆေးဆိုင် ကို ▭ ဆေး ဝယ် ။ ↓ ↓ ↙ ↙ ↙ ဆေးဆိုင် အဲ့ဒီမှာ ဆေး ဝယ် ။	i=5, j=9, k=0 Z=9

Evaluation Methodology

- Automatic evaluation of machine translation (MT) quality is essential to developing high-quality machine translation systems because human evaluation is time consuming, expensive, and irreproducible.
- In NLP, there are many kinds of automatic evaluation methodology (e.g., **BLEU**, NIST, PER, TER, **WER**, MERT and **RIBES,** etc.).
- **BLEU** measures how many words overlap in a given translation when compared to a reference translation.
- **RIBES** used rank correlation coefficients based on word order to compare SMT and reference translations and it overcomes reordering.

BLEU –BiLingual Evaluation Understudy

- It computes precision for n-grams of size 1 to 4.
- It ranges from 0-100, the higher the score, the more the translation correlates to a human translation.

$$\mathrm{BP} = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases}.$$

Then,

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

where, c -Hypothesis or Candidate Sentence Length,
r -Reference Sentence Length
BP -Brevity Penalty,
p -Precision

BLEU –BiLingual Evaluation Understudy

- Hypothesis: ကျွန်တော့် အစ်ကို ပျောက် သွား တယ် ကို ရှာ ပေး ပါ ။
- Reference: ကျွန်တော့် အစ်ကို ကို ရှာ ပေး ပါ ။
- c = 10, r = 7 and it satisfies c > r so BP = 1
- Use **N=4** and uniform weight $\boldsymbol{w_n = \frac{1}{N}}$

N-gram	W_n	P_n	Log P_n	W_n * log P_n
1	1/4	7/10	-0.1549	-0.0387
2	1/4	5/9	-0.2553	-0.0638
3	1/4	3/8	-0.426	-0.1065
4	1/4	2/7	-0.5441	-0.136
Total				-0.345
BLEU = 1 * exp (-0.345) = **0.7082 = 70.82%**				

RIBES –Rank-based Intuitive Bilingual Evaluation Score

- Kendall's τ method is used to calculate RIBES score because of a smaller error sensitivity or more robust and more efficient.
- Kendall's τ measures the direction of differences in rank.

$$\text{Kendall's } \tau = 2 * \frac{numbers\ of\ increasing\ pairs}{numbers\ of\ all\ pairs} - 1$$

- Kendall's τ has the **range [−1, 1**].
- The rank can be normalized to ensure positive values.
  - Normalized Kendall's τ (NKT) = (τ + 1)/2

RIBES –Rank-based Intuitive Bilingual Evaluation Score

- R0 : ကျွန်တော့် အစ်မ လှေကား ပေါက လိမ့်ကျ ဆေးရုံတင် ။
- H0 : ငှါ လှေကား ပေါက ချော်ကျ အစ်မ သွား ။

By removing non-aligned words by one-to-one correspondence,

- R1 : ကျွန်တော့်$_1$ အစ်မ$_2$ လှေကား$_3$ ပေါက$_4$ လိမ့်ကျ$_5$ ။$_6$
- H1 : ငှါ$_1$ လှေကား$_3$ ပေါက$_4$ ချော်ကျ$_5$ အစ်မ$_2$ ။$_6$
- Word order of R1 : [1, 2, 3, 4, 5, 6]
- Word order of H1 : [1, 3, 4, 5, 2, 6]
- Number of increasing pairs =12
- Number of all pairs = 15
- Kendall's $\tau = 2 * \frac{12}{15} - 1 = 0.6$
- Normalized Kendall's τ (NKT) = (0.6+ 1)/2 = **0.8 = 80%**

Experiments

1. Corpus Statistic

- The experiment uses a parallel corpus for Myanmar written text, Myanmar sign language and Myanmar SignWriting in Emergency domain.
- Current parallel corpus size is limited and 10-fold cross validation is done for all experiments.
- Total sentences in the corpus contain 888 sentences.
  - 600 sentences are used for training.
  - 138 sentences are used for development.
  - 150 sentences are used for testing.

Figure 5: Cross Validation Strategy

Experiments

2. Moses SMT system

- It uses PBSMT, HBPSMT, OSM system provided by the Moses toolkit.
- Word segmented source language are aligned with the word segmented target language using GIZA++.
- KENLM is used as a language model.

Translation between MWT and MSW

src-trg	MWT (word)-MSW(word)		
	PBSMT	**HPBSMT**	**OSM**
my-sw	**12.476** (0.6598)	12.328 (0.6525)	12.427 **(0.6426)**
sw-my	22.268 (0.7184)	22.209 (0.7281)	**23.063 (0.7295)**

Table 1: BLEU and RIBES scores of word-word segmentation pair for PBSMT, HPBSM and OSM

Translation between MWT and MSW

src-trg	MWT (syllable)-MSW(word)		
	PBSMT	**HPBSMT**	**OSM**
my-sw	9.108 (0.6358)	**9.432 (0.6388)**	9.273 (0.6289)
sw-my	20.345 (0.7231)	20.081 (0.7234)	**21.628 (0.7288)**

Table 2: BLEU and RIBES scores of syllable-word segmentation pair for PBSMT, HPBSM and OSM

Translation between MWT and MSL

src-trg	MWT (word)-MSL(word)		
	PBSMT	**HPBSMT**	**OSM**
my-sl	27.359 (0.7436)	**28.825 (0.7591)**	27.376 (0.7398)
sl-my	29.1 (0.7772)	29.1 (0.7836)	**30.604 (0.7675)**

Table 3: BLEU and RIBES scores of word-word segmentation pair for PBSMT, HPBSM and OSM

Translation between MWT and MSL

src-trg	MWT (syllable)-MSL(syllable)		
	PBSMT	**HPBSMT**	**OSM**
my-sl	31.877 (0.8348)	**32.19 (0.8352)**	32.125 (0.8342)
sl-my	28.879 (0.8384)	30.414 (**0.8421**)	**30.767** (0.8391)

Table 4: BLEU and RIBES scores of syllable-syllable segmentation pair for PBSMT, HPBSM and OSM

Translation between MSL and MSW

src-trg	MSL (word)-MSW(word)		
	PBSMT	**HPBSMT**	**OSM**
sl-sw	**41.122 (0.8577)**	40.82 (0.8553)	40.823 (0.8567)
sw-sl	**50.184** (0.8759)	50.073 **(0.8764)**	49.497 (0.8741)

Table 5: BLEU and RIBES scores of word-word segmentation pair for PBSMT, HPBSM and OSM

Translation between MSL and MSW

src-trg	MSL (syllable)-MSW(word)		
	PBSMT	**HPBSMT**	**OSM**
sl-sw	32.094 (0.8298)	**33.069 (0.8314)**	31.554 (0.8274)
sw-sl	47.674 (0.8674)	47.703 (0.8692)	**48.125 (0.8695)**

Table 6: BLEU and RIBES scores of syllable-word segmentation pair for PBSMT, HPBSM and OSM

Error Analysis

- **Word Error Rate –WER**
- It is dynamic programming to find an optimal alignment between the hypothesis of machine translation and the reference translation
- Formula of WER: WER= $\frac{S+I+D}{N}$
  - N = number of words
  - S = number of substituted words
  - I = number of inserted words
  - D = number of deleted words

Error Analysis

- Ref : ငါ အရေးပေါ် လူနာတင် ကား လိုချင် လို့ အမြန် ခေါ် ပါ ။
- Hyp : ငါ အရေးပေါ် လူနာတင် ကား အမြန် **** ခေါ်ပေး လိုချင် **** ။

WER errors:

Reference	**Hypothesis**	**Error Type**
လိုချင်	အမြန်	Substitution
လို့		Deletion
အမြန်	ခေါ်ပေး	Substitution
ခေါ်	လိုချင်	Substitution
ပါ		Deletion

- In this case, **S=3, D=2, C=5, N=10** for PBSMT and its **WER** is equal to **50%.**

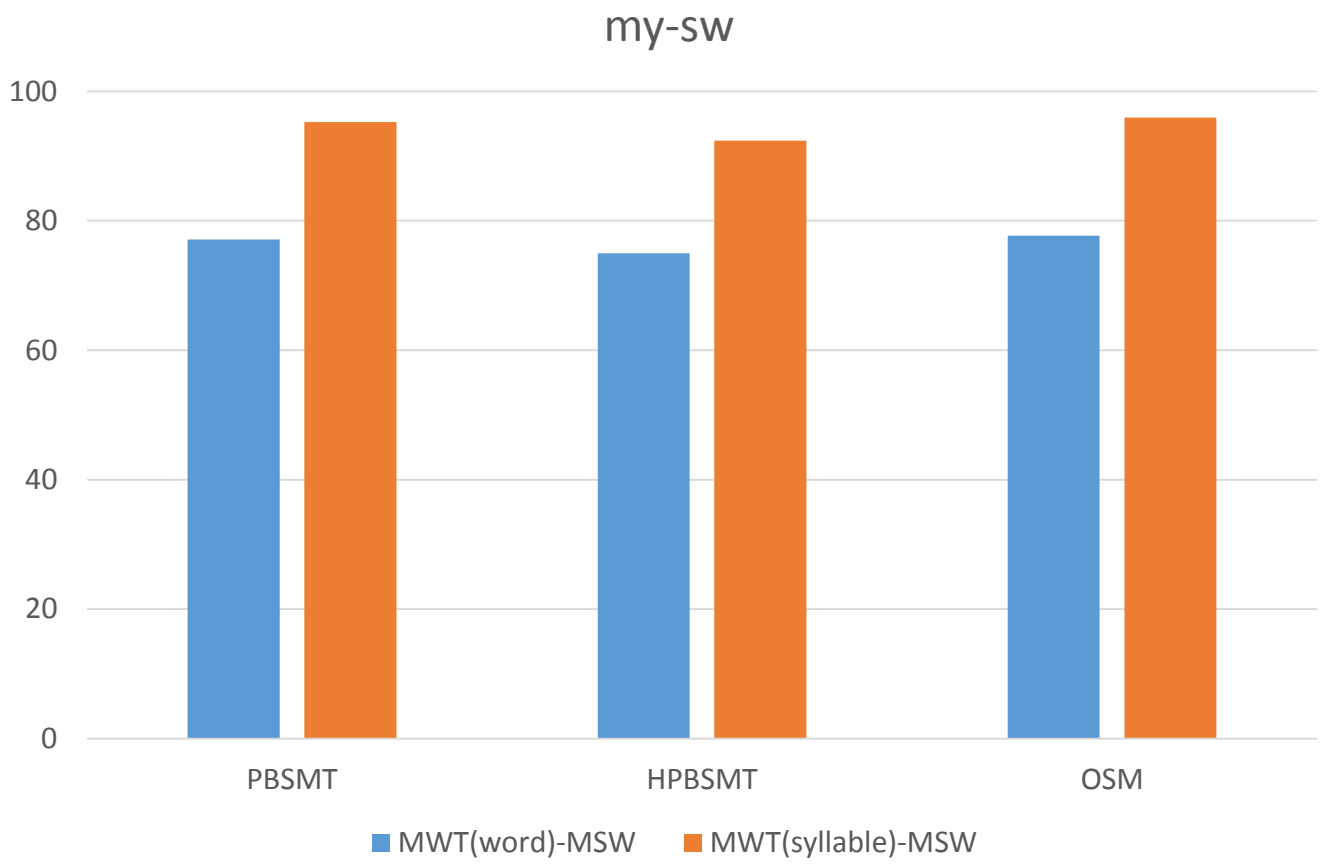


Figure 6 : WER of machine translation from Myanmar written text to Myanmar SignWriting

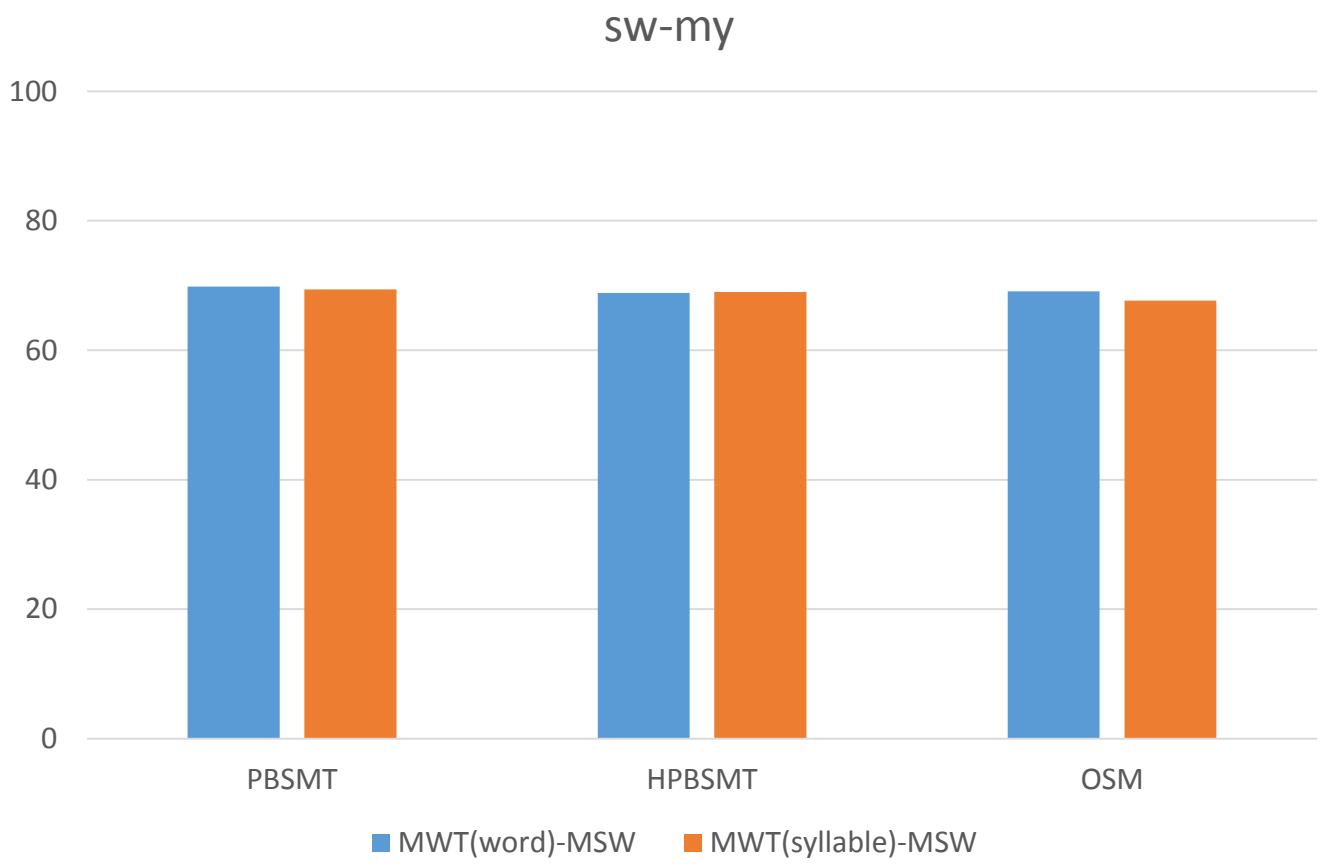


Figure 7 : WER of machine translation from Myanmar written text to Myanmar sign language

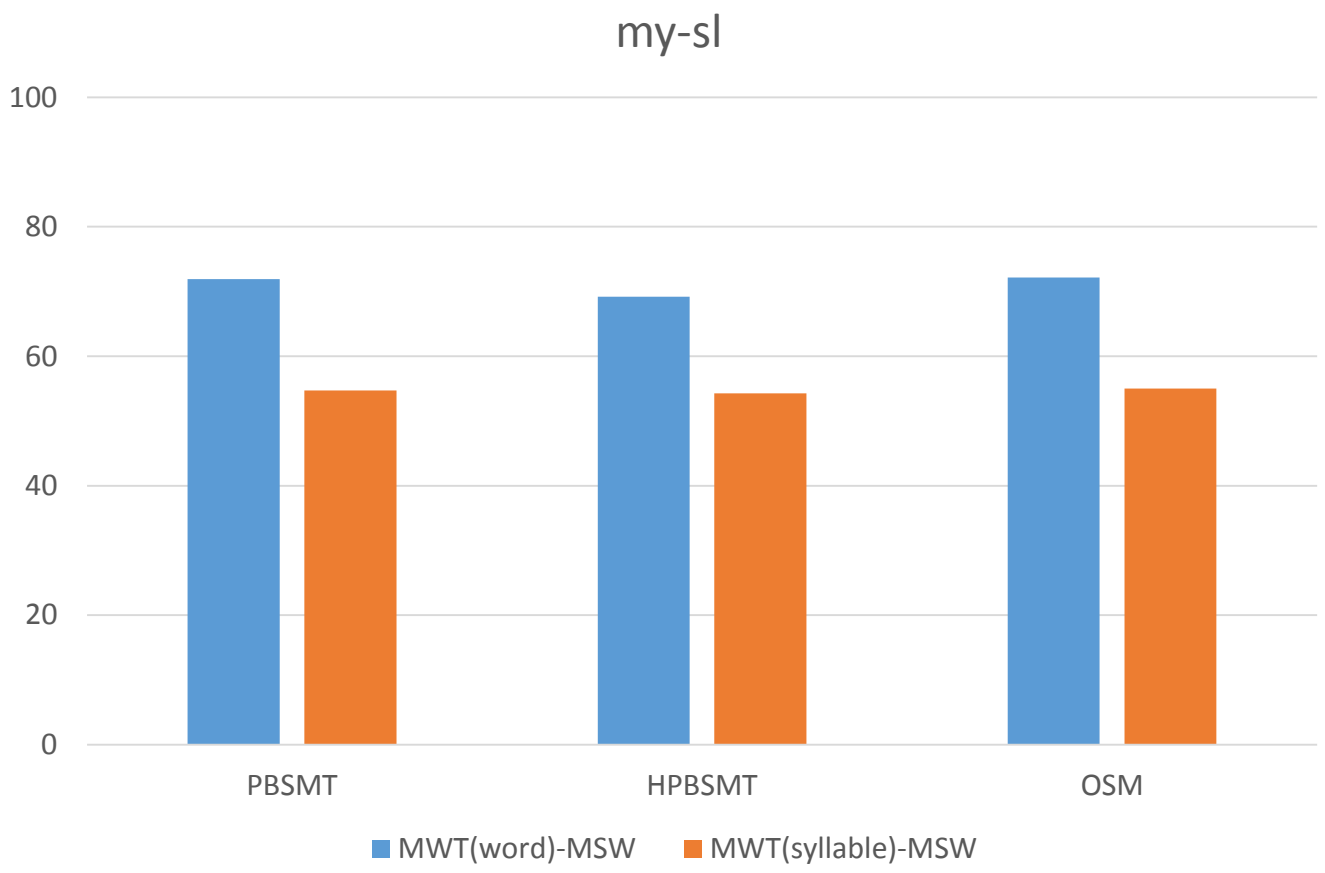


Figure 8 : WER of machine translation from Myanmar written text to Myanmar sign language

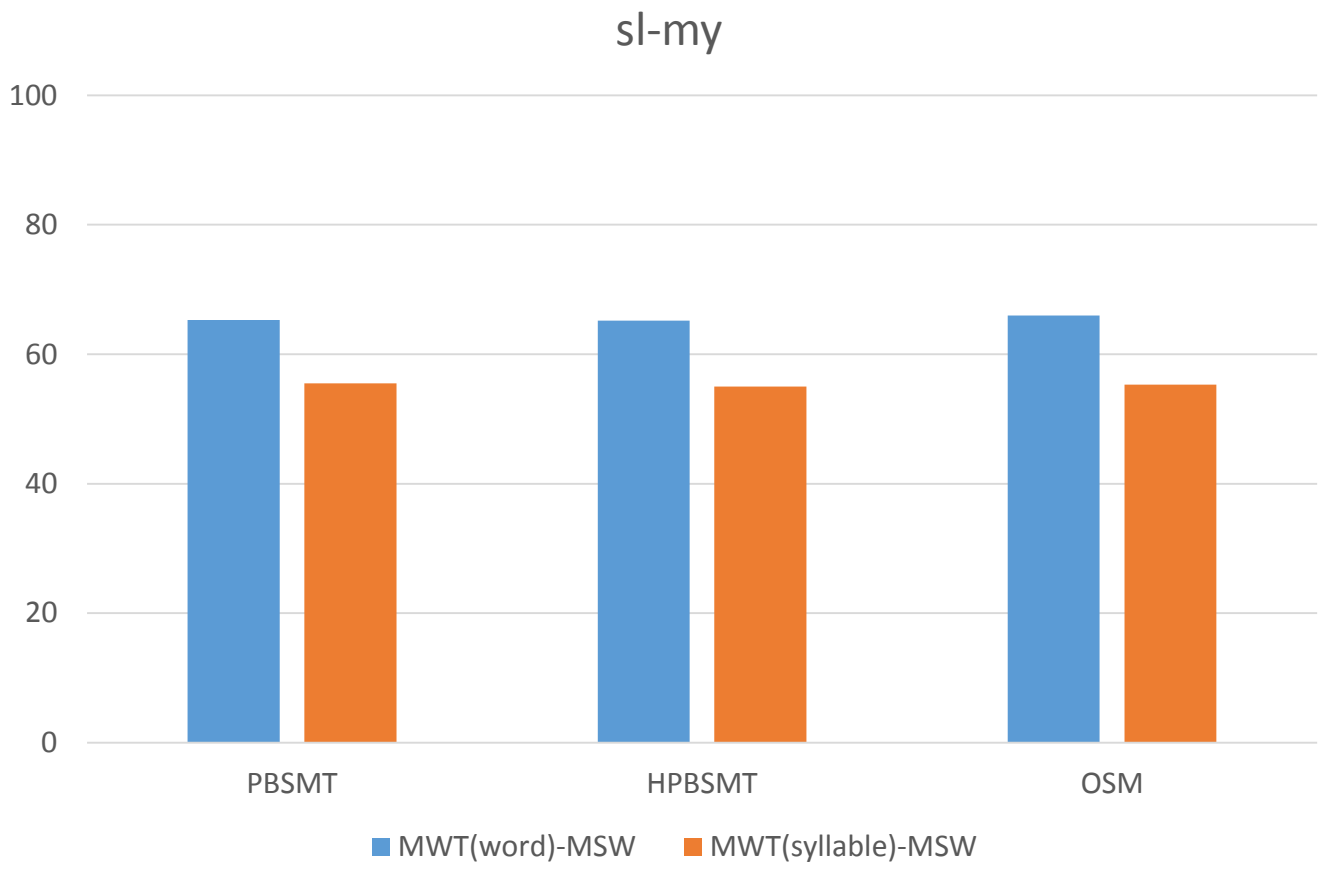


Figure 9 : WER of machine translation from Myanmar sign language to Myanmar written text

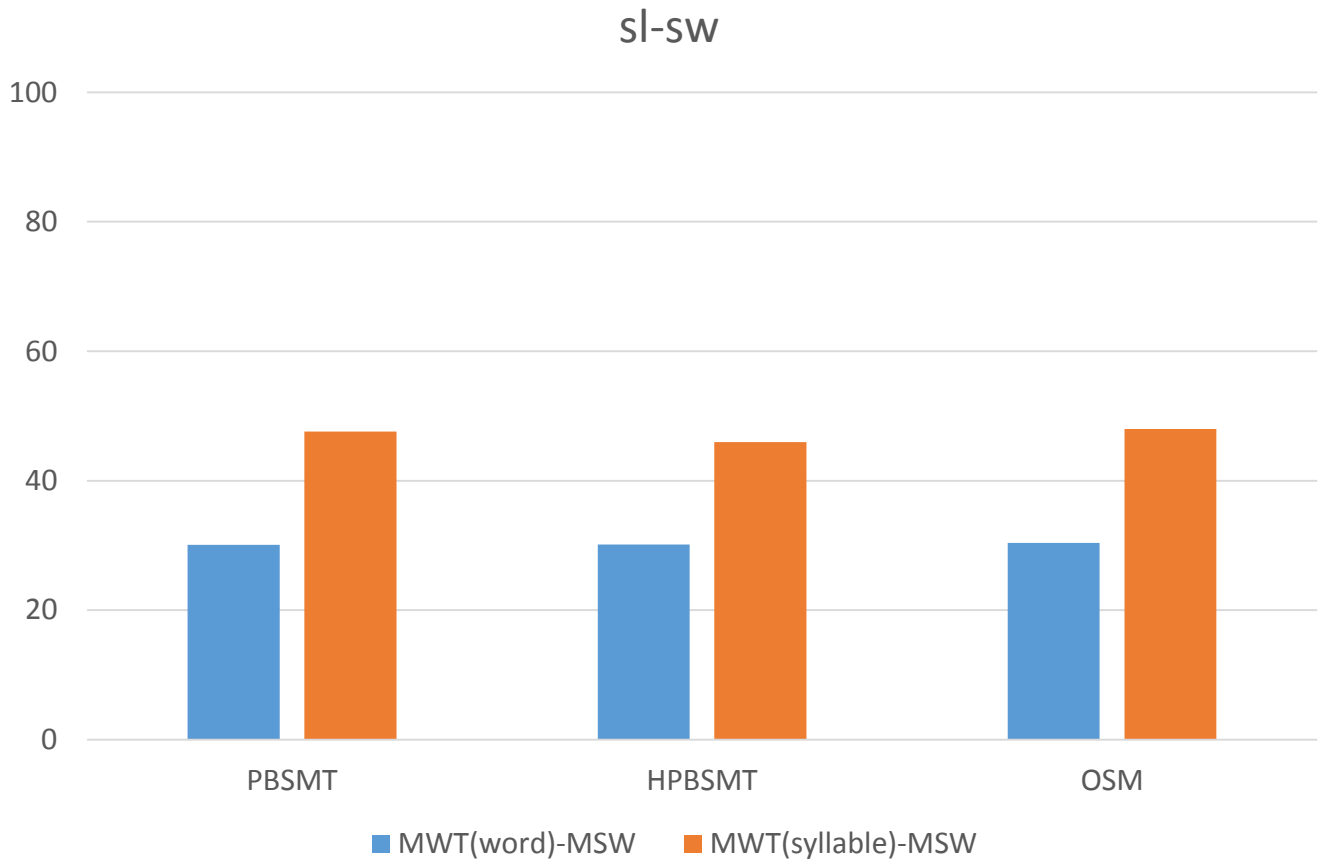


Figure 10 : WER of machine translation from Myanmar sign language to Myanmar SignWriting

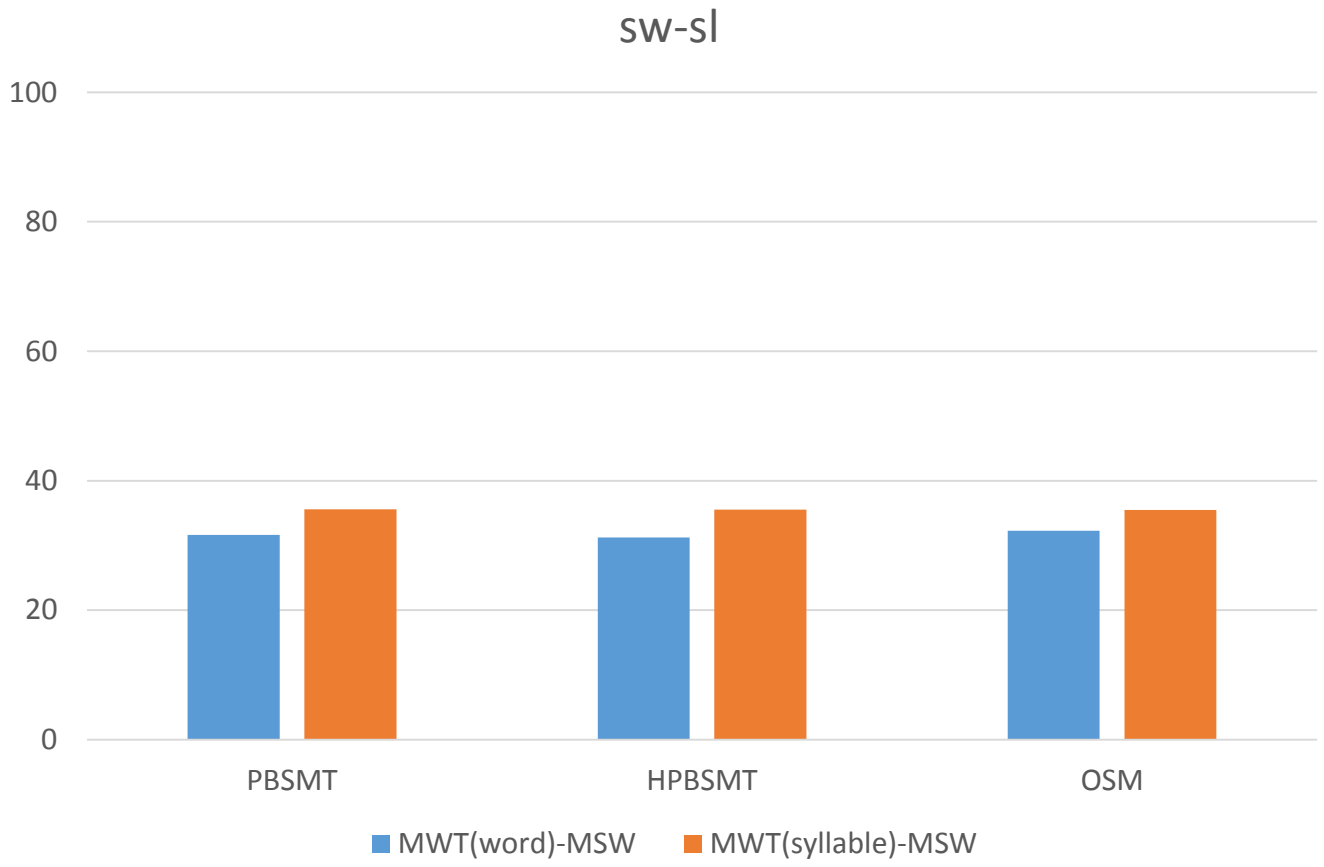


Figure 11 : WER of machine translation from Myanmar SignWriting to Myanmar sign language

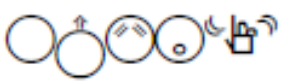
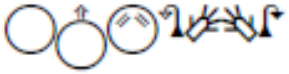
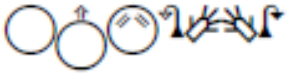
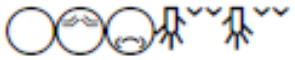
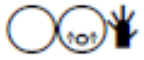
No	References	Hypothesis	Ref-Hyp Description in Myanmar Language	Freq
1			ဆရာဝန် -> ဆရာဝန်	3
2			ဘယ်လောက်လဲ -> ဘာလဲ	3
3				3
4			မသုံးနဲ့ -> မလုပ်နဲ့	3
5			မလုပ်နဲ့ -> မလုပ်နဲ့	3
6			ဘယ်သူ -> ဘာလဲ	2
7			ကယ်ဆယ် -> ကျေးဇူးပြု၍	2
8				2
9			ညစ်ပတ် -> ရေ	2
10			ရှိ -> မသေချာ	1

Table 7: Top 10 confusion pairs of OSM model in MWT-MSW translation

Error Analysis

- From detail analysis on confusion pairs of three SMT approaches, most of them are caused by four main reasons as follow:
  - **The nature of sign language,**
  - **Alignment error between source and target,**
  - **Some errors in the references or human mistakes,**
  - **Limited size of the training data**.

Conclusion

- The research introduced the first evaluation of machine translation between Myanmar Written Text and Myanmar SignWriting using three SMT approaches.
- Translation performances are compared and error rate is analyzed.
- From the overall experiment results, word segmentation achieved better performance than syllable segmentation in both MWT-MSW and MSL-MSW because of the grammar structure of SignWriting nature.
- The difference is that syllable segmentation achieved better performance than word segmentation in MWT-MSL pair.
- It helps to overcome the communication problem between Deaf and hearing people by introducing the SignWriting.

Future Works

- The future work will 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.
- The performance analysis can do with current SMT approaches and other machine translation approaches.

List of Publications

- Swe Zin Moe, Hlaing Myat Nwe, **Hnin Wai Wai Hlaing**, Ye Kyaw Thu, Hnin Aye Thant, Nandar Win Min, “Myanmar Sign Language (MSL) Corpus for Emergency Domain”, PACLING2017 conference, Yangon, Myanmar. (Demo and Poster)
- Swe Zin Moe, Ye Kyaw Thu, Hlaing Myat Nwe, **Hnin Wai Wai Hlaing**, Ni Htwe Aung, Hnin Aye Thant, Nandar Win Min, “Corpus Building for Machine Translation between Myanmar Sign Language and Myanmar Written Text”, World Deaf Day 2017, 14th Sept. 2017, Mandalay Community Center, Chan Aye Tharzan Township, Mandalay, Myanmar. (Poster)
- Swe Zin Moe, **Hnin Wai Wai Hlaing**, Ye Kyaw Thu, Hlaing Myat Nwe, Ni Htwe Aung, Hnin Aye Thant, Nandar Win Min, “မြန်မာ လက်သင်္ကေတပြဘာသာစကားမှ မြန်မာစကားပြောစာကြောင်းသို့ ကွန်ပျူတာသုံး ဘာသာပြန် သုတေသန”, International Day of Persons with Disabilities 2017, 3rd Dec. 2017, Wilson Hotel, No.31(E), Yangon-Mandalay Main Road, Maha Aung Myay Township, Mandalay, Myanmar. (Demo and Poster)

List of Publications

- Hlaing Myat Nwe, Ye Kyaw Thu, **Hnin Wai Wai Hlaing**, Swe Zin Moe, Ni Htwe Aung, Hnin Aye Thant, Nandar Win Min, "Two Fingerspelling Keyboard layouts for Myanmar SignWriting", International Day of Persons with Disabilities 2017, 3rd Dec. 2017, Wilson Hotel, No.31(E), Yangon-Mandalay Main Road, Maha Aung Myay Township, Mandalay, Myanmar. (Demo and Poster)
- Hlaing Myat Nwe, Ye Kyaw Thu, **Hnin Wai Wai Hlaing**, Swe Zin Moe, Ni Htwe Aung, Hnin Aye Thant, Nanda Win Min, "Two Fingerspelling Keyboard Layouts for Myanmar SignWriting", In Proceedings of ICCA2018, February 22-23, 2018, Yangon, Myanmar, pp. 290-298. (Paper)
- Swe Zin Moe, Ye Kyaw Thu, **Hnin Wai Wai Hlaing**, Hlaing Myat Nwe, Ni Htwe Aung, Hnin Aye Thant, Nandar Win Min, "Statistical Machine Translation between Myanmar Sign Language and Myanmar Written Text", In Proceedings of ICCA2018, February 22-23, 2018, Yangon, Myanmar, pp. 217-227. (Paper)
- **Hnin Wai Wai Hlaing**, Ye Kyaw Thu, Swe Zin Moe, Hlaing Myat Nwe, Ni Htwe Aung, Nandar Win Min, Hnin Aye Thant, “Statistical Machine Translation between Myanmar Sign Language and Myanmar SignWriting”, at the First IEEE International Symposium on Artificial Intelligence for ASEAN Development, ASEAN-AI2018, Phuket, Thailand, 26th March 2018. (Paper)

References

- Philipp Koehn., “Statistical Machine Translation: the basic, the novel, and the speculative”
- Nadir Durrani, Helmut Schmid, Alexander Fraser, Philipp Koehn, Hinrich Schütze., “The Operation Sequence Model –Combining N-Gram-based and Phrase-based Statistical Machine Translation”
- Yoav Artzi., “Phrase-based Translation”
- Jörg Tiedemann. Hierarchical phrase-based machine translation
- Hideki Isozaki, Tsutomu Hirao, Kevin Duh, Katsuhito Sudoh, Hajime Tsukada.,” Automatic Evaluation of Translation Quality for Distant Language Pairs”
- Ameera M.Almasoud and Hend S. Al-Khalifa., “A Proposed Semantic Machine Translation System for translating Arabic text to Arabic sign language”
- Win Pa Pa, Ye Kyaw Thu, Andrew Finch, Eiichiro Sumita,. “A Study of Statistical Machine Translation Methods for Under Resourced Languages”
- Phillip Koehn,. “Statistical Machine Transaltion”
- The population and housing census of Myanmar,2014
- Valerie Sutton; International SignWriting Alphabet (2010)
- Valerie Sutton and Adam Frost; SignWriting Hand Symbols in ISWA2010: Manual 2
- https://github.com/ye-kyaw-thu/sylbreak
- https://github.com/ye-kyaw-thu/MSL4Emergency
- Martin Thoma.: Word Error Rate Calculation, (2013) https://martin-thoma.com/word-error-rate-calculation/

Thank You for your kind attention…