Statistical Machine Translation between Kachin and Rawang

1st Ye Kyaw Thu
Language and Semantic Technology Research Team (LST)
National Electronics and Computer Technology Center (NECTEC)
PathumThani, Thailand
ka2pluskha2@gmail.com

2nd Manar Hti Seng
Faculty of Computer Science
University of Computer Studies (Banmaw)
Banmaw, Myanmar
junjasmine298@gmail.com

3rd Thazin Myint Oo
Natural Language Process Lab.
University of Computer Studies, Yangon (UCSY)
Yangon, Myanmar
thazinmyintoo@ucsy.edu.mm

4th Dee Wom
Faculty of Computer Science
University of Computer Studies (Banmaw)
Banmaw, Myanmar
Mgnyochaw7@gmail.com

5th Hpau Myang Thint Nu
Faculty of Computer Science
University of Computer Studies (Banmaw)
Banmaw, Myanmar
thintnuhpaumyang@gmail.com

6th Seng Mai
Faculty of Computer Science
University of Computer Studies (Banmaw)
Banmaw, Myanmar
sarinaryuri@gmail.com

7th Thepchai Supnithi
Language and Semantic Technology Research Team (LST)
National Electronics and Computer Technology Center (NECTEC)
PathumThani, Thailand
thepchai@gmail.com

8th Khin Mar Soe
Natural Language Process Lab.
University of Computer Studies, Yangon (UCSY)
Yangon, Myanmar
khinmarsoe@ucsy.edu.mm

***Abstract*—This paper contributes the first evaluation of the quality of machine translation between Kachin and Rawang. We also developed a Kachin-Rawang parallel corpus (around 10K sentences) based on the Myanmar language of ASEAN MT corpus. The 10 folds cross-validation experiments were carried out using three different statistical machine translation approaches: phrase-based, hierarchical phrase-based, and the operation sequence model (OSM). The results show that all three statistical machine translation approaches give higher and comparable BLEU and RIBES scores for both Kachin to Rawang and Rawang to Kachin machine translations. OSM approach achieved the highest BLEU and RIBES scores among three approaches machine translation.**

***Index Terms*—Statistical Machine Translation, Under-resourced languages, Dialects, Kachin, Rawang**

I. Introduction

Our main motivation for this research is to investigate SMT performance for Kachin and Rawang language pair. The Kachin language is closely related to Rawang language and it is often considered as dialect of Kachin language. The state-of-the-art techniques of statistical machine translation (SMT) [1], [2] demonstrate good performance on translation of languages with relatively similar word orders [3]. To date, there have been some studies on the SMT of Myanmar language. Ye Kyaw Thu et al. (2016) [4] presented the first large-scale study of the translation of the Myanmar language. A total of 40 language pairs were used in the study that included languages both similar and fundamentally different from Myanmar. The results show that the hierarchical phrase-based SMT (HPBSMT) [5] approach gave the highest translation quality in terms of both the BLEU [6] and RIBES scores [7]. Win Pa Pa et al (2016) [8] presented the first comparative study of five major machine translation approaches applied to low-resource languages. PBSMT, HPBSMT, tree-to-string (T2S), string-to-tree (S2T) and OSM translation methods to the translation of limited quantities of travel domain data between English and Thai, Laos, Myanmar in both directions. The experimental results indicate that in terms of adequacy (as measured by BLEU score), the PBSMT approach produced the highest quality translations. Here, the annotated tree is used only for English language for S2T and T2S experiments. This is because there is no publicly available tree parser for Lao, Myanmar and Thai languages. According to our knowledge, there is no publicly available tree parser for both Kachin and Rawang languages and thus we cannot apply S2T and T2S approaches for Kachin-Rawang language pair. From their RIBES scores, we noticed that OSM approach achieved best machine translation performance for Myanmar to English translation. Moreover, we learned that OSM approach gave highest translation performance between Khmer (the official language of Cambodia) and twenty other languages, in both directions [9].

Relating to Myanmar langauge dialects, Thazin Myint Oo et al. (2018) [25] contributed the first PBSMT, HPBSMT and OSM machine translation evaluations between Myanmar and Rakhine. The experiment was used the 18K Myanmar-Rakhine parallel corpus that constructed to analyze the behavior of a dialectal Myanmar-Rakhine machine translation. The results showed that higher BLEU

(57.88 for Myanmar-Rakhine and 60.86 for Rakhine-Myanmar) and RIBES (0.9085 for Myanmar-Rakhine and 0.9239 for Rakhine-Myanmar) scores can be achieved for Rakhine-Myanmar language pair even with the limited data. Based on the experimental results of previous works, in this paper, the machine translation experiments between Kachin and Rawang were carried out using PBSMT, HPBSMT and OSM.

II. Related Work

Karima Meftouh et al. built PADIC (Parallel Arabic Dialect Corpus) corpus from scratch, then conducted experiments on cross dialect Arabic machine translation [10]. PADIC is composed of dialects from both the Maghreb and the Middle-East. Some interesting results were achieved even with the limited corpora of 6,400 parallel sentences. Using SMT for dialectal varieties usually suffers from data sparsity, but combining word-level and character-level models can yield good results even with small training data by exploiting the relative proximity between the two varieties [11]. Friedrich Neubarth et al. described a specific problem and its solution, arising with the translation between standard Austrian German and Viennese dialect. They used hybrid approach of rule-based preprocessing and PBSMT for getting better performance. Pierre-Edouard Honnet et al. proposed solutions for the machine translation of a family of dialects, Swiss German, for which parallel corpora are scarce [12]. They presented three strategies for normalizing Swiss German input in order to address the regional and spelling diversity. The results show that character-based neural MT was the most promising one for text normalization and that in combination with PBSMT achieved 36% BLEU score.

III. Kachin State and Kachin People

The Kachin State is situated in north Myanmar and is the place where the most of jinghpaw peoples live in. It lies between north latitude 23°27’and 28°25’ longitude 96°0’ and 98°44’. The area of Kachin State is 89,041km (34,379 sq mi). The capital of Kachin State is Myitkyina and Bhamo is the second largest historic city of jinghpaw peoples. About 2 millions of jinghpaw peoples live in Myanmar. In jinghpaw, there are two groups called Bhamo jinghpaw and Myitkyina jinghpaw. Jinghpaw peoples also live in Shan State. Jinghpaw peoples are on of the ethnic groups in Myanmar. Jinghpaw comprises six tribes or subdivisions: Lisu, Lashi, Rawang, Zaiwa, Lhao Vo. All have their own language and literature. As they all comes from Jinghpaw, they can also speak or use the Jinghpaw language. Jinghpaw alphabet is based on Lathin script. The Jinghpaw literature was strated using in the era of “King Min Done Min” (1853-1878). However, the literature that Kachin peoples are using now was written on May 5,1895 by Dr. Ola Hanson, in the era of “King Thi Paul” (1878-1885) .

IV. Kachin and Rawang Languages

A. Kachin or Jingpho Language

Jingpho (Jinghpaw, Chingp’o) or Kachin (Burmese: ကချင်ဘာသာ [kət` bàðà]) is a Tibeto-Burman language of the Sal branch mainly spoken in Kachin State, Burma and Yunnan, China. There are a lot of meanings for Jingpho. In the Jingpho language, Jingpho means people. The term ”Kachin language” can refer either to the Jingpho language or to a group of languages spoken by various ethnic groups in the same region as Jingpo: Lisu, Lashi, Rawang, Zaiwa, Lhao Vo, Achang and Jingpho. These languages are from distinct branches of the highest level of the Tibeto-Burman family. The Jingpho alphabet is based on the Latin script. The ethnic Jingpho (or Kachin) are the primary speakers of Jingpho language, numbering approximately 900,000 speakers. The Turung of Assam in India speak a Jingpho dialect with many Assamese loanwords, called Singpho.

B. Rawang Language

Rawang peoples is the sub-group of Kachin (Jinghpaw) in Myanmar. Rawang peoples live in northern Kachin state: Puato, Machanbaw, Naungmaw, Kawnglangphu, and Pannandin townships. 70,000 of Rawang people live in Myanmar. There are four enthic roup in Rawang. They are Lungmi, Matwang, Daru and Tangsar. The Normal (-) , High (´) and Low (‘) symbols:

C. Grammar of Myanmar, Kachin and Rawang

Kachin and Rawang sentences use several punctuation symbols “.”, “,”, “?”, ““” , “””, “‘”, “’”,“!”, “:”, “;” and “-”. Myanmar (my), Kachin (kc) and Rawang (rw) languages have same word order (i.e. Subject, Object and Verb). Some example of parallel sentences in Myanmar(my), Kachin(kc) and Rawang(rw) are as follows :

my: နေကောင်းလား ။
kc: Hkam kaja ai i .
rw: PÀMVRÀ ÍÈ MÁ .

my: လုံချည် တစ်ထည် ဘယ်လောက်လဲ ။
kc: Ba hkang langai kade rai ?
rw: SHVRØM TÌQ DUNG KÀDVNGTĒ ÍÈ lĒ .

my: ထမင်းစား ပြီးပြီ လား ။
kc: Shat sha sai i ?
rw: VMPÀ VM BØĪ MÁ .

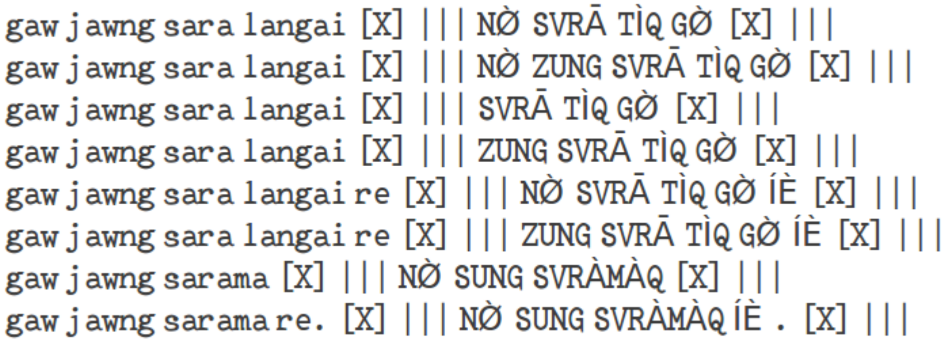
my: ကလေးများ ကစား နေကြသည် ။
kc: Ma ni gasup taw nga ma ai .
rw: CVMRE RĪ GVSØP MĒ .

V. Methodology

In this section, we describe the methodology used in the machine translation experiments for this paper.

A. Phrase-Based Statistical Machine Translation

A PBSMT translation model is based on phrasal units [1]. 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 [14]. The phrase translation model is based on noisy channel model. To find best translation $\hat{e}$ that maximizes the translation probability $\mathbf{P}(f)$ given the source sentences; mathematically. Here, the source language is French and the target language is an English. The translation of a French sentence into an English sentence is modeled as equation 1.



```
gaw jawng sara langai [X] ||| NÒ SVRĀ TÌQ GÒ [X] |||
gaw jawng sara langai [X] ||| NÒ ZUNG SVRĀ TÌQ GÒ [X] |||
gaw jawng sara langai [X] ||| SVRĀ TÌQ GÒ [X] |||
gaw jawng sara langai [X] ||| ZUNG SVRĀ TÌQ GÒ [X] |||
gaw jawng sara langai re [X] ||| NÒ SVRĀ TÌQ GÒ ÍÈ [X] |||
gaw jawng sara langai re [X] ||| ZUNG SVRĀ TÌQ GÒ ÍÈ [X] |||
gaw jawng sarama [X] ||| NÒ SUNG SVRÀMÀQ [X] |||
gaw jawng sarama re. [X] ||| NÒ SUNG SVRÀMÀQ ÍÈ . [X] |||
```

Fig. 1: Some examples of hierarchical phrase-based grammar between Kachin and Rawang phrases

$$\hat{e} = argmax_e \mathbf{P}(e|f) \tag{1}$$

Applying the Bayes' rule, we can factorized into three parts.

$$P(e|f) = \frac{\mathbf{P}(e)}{\mathbf{P}(f)} \mathbf{P}(f|e) \tag{2}$$

The final mathematical formulation of phrase-based model is as follows:

$$argmax_e \mathbf{P}(e|f) = argmax_e \mathbf{P}(f|e) \mathbf{P}(e) \tag{3}$$

We note that denominator $\mathbf{P}(f)$ can be dropped because for all translations the probability of the source sentence remains the same. The $\mathbf{P}(e|f)$ variable can be viewed as the bilingual dictionary with probabilities attached to each entry to the dictionary (phrase table). The $\mathbf{P}(e)$ variable governs the grammaticality of the translation and we model it using n-gram language model under the PBMT paradigm.

B. *Hierarchical Phrase-Based Statistical Machine Translation*

The hierarchical phrase-based SMT approach is a model based on synchronous context-free grammar [14]. 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 re-ordering process. The re-ordering is represented explicitly rather than encoded into a lexicalized re-ordering model (commonly used in purely phrase-based approaches). This makes the approach particularly applicable to language pairs that require long-distance re-ordering during the translation process [15]. Some examples of hierarchical phrase based grammar between Dawei and Myanmar phrases are shown in Figure 1.

C. *Operation Sequence Model*

The operation sequence model which combines the benefits of two state-of-the-art SMT frameworks named n-gram-based SMT and phrase-based SMT. This model simultaneously generate source and target units and does not have spurious ambiguity that is based on minimal translation units [16] [17]. It is a bilingual language model that also integrates reordering information. OSM motivates better reordering mechanism that uniformly handles local and non-local reordering and strong coupling of lexical generation and reordering. It means that OSM can handle both short and long distance reordering. The operation types are such as generate, insert gap, jump back and jump forward which perform the actual reordering. The following shows an example translation process of English sentence "Please sit here" into Myanmar language with the OSM.

Source: Please sit here
Target: ကျေးဇူးပြုပြီး ဒီမှာ ထိုင်

Operation 1: Generate (Please, ကျေးဇူးပြုပြီး)
Operation 2: Insert Gap
Operation 3: Generate (here, ကျေးဇူးပြုပြီး ဒီမှာ)
Operation 4: Jump Back (1)
Operation 5: Generate (sit, ကျေးဇူးပြုပြီး ဒီမှာ ထိုင်)

VI. Experiment

A. *Corpus Statistics*

We used 10K Myanmar sentences (without name entity tags) of the ASEAN-MT Parallel Corpus [18], which is a parallel corpus in the travel domain. It contains six main categories and they are people (greeting, introduction and communication), survival (transportation, accommodation and finance), food (food, beverage and restaurant), fun (recreation, traveling, shopping and nightlife), resource (number, time and accuracy), special needs (emergency and health). Manual translation to Kachin and Rawang was done manually. We held 10-fold cross-validation experiments and used 8,468 to 8,519 sentences for training, 500 sentences for development and 985 to 1,026 sentences for evaluation respectively.

B. *Moses SMT System*

We used the PBSMT, HPBSMT and OSM system provided by the Moses toolkit [19] for training the PBSMT, HPBSMT and OSM statistical machine translation systems. The word segmented source language was aligned with the word segmented target language using GIZA++ [20]. The alignment was symmetrized by grow-diag-final and heuristic [1]. The lexicalized reordering model was trained with the msd-bidirectional-fe option [21]. We use KenLM [22] for training the 5-gram language model with modified Kneser-Ney discounting [23]. Minimum error rate training (MERT) [24] was used to tune the decoder parameters and the decoding was done using the Moses decoder (version 2.1.1). We used default settings of Moses for all experiments.

VII. 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) [6] and the other was the Rank-based Intuitive Bilingual Evaluation Measure (RIBES) [7]. The BLEU score measures the precision of n-gram (over all n 4 in our case) with respect to a reference translation with a penalty for short translations [6]. Intuitively, the BLEU score measures the adequacy of the translation and large BLEU scores are better. RIBES is an automatic evaluation metric based on rank correlation coefficients modified with precision and special care is paid to word order of the translation results. The RIBES score is suitable for distance language pairs such as Myanmar and English. Large RIBES scores are better.

VIII. Results and Discussion

The BLEU and RIBES score results for machine translation experiments with PBSMT, HPBSMT and OSM are shown in Table 1. Bold numbers indicate the highest scores among three SMT approaches. The RIBES scores are inside the round brackets. Here, “kc” stands for Kachin, “rw” stands for Rawang, “src” stands for source language and “tgt” stands for target language respectively.

From the results, OSM method achieved the highest BLEU and RIBES score for both Kachin to Rawang and Rawang to Kachin bi-directional machine translations. Interestingly, the BLEU and RIBES score of all three methods are comparable performance. Our results with current parallel corpus indicate that Rawang to Kachin machine translation is better performance (around 3 BLEU and 0.02 RIBES scores higher) than Kachin to Rawang machine translation direction.

As we expected, generally, machine translation performance of all three SMT approaches between Kachin and Rawang languages achieved good scores for both BLEU and RIBES. The reason is that as we mentioned in Section IV-A and Section IV-B, the two languages, Kachin and Rawang are close languages. We assume that long distance reordering is relatively rare and only local reordering is enough for the Kachin-Rawang language pair.

IX. Error Analysis

We also used the SCLITE (score speech recognition system output) program from the NIST scoring toolkit SCTK version 2.4.10 [26] for making dynamic programming based alignments between reference and hypothesis strings for detail analysis on translation errors (WER: Word Error Rate). The formula for WER can be stated as equation 4:

$$WER = (I + D + S)100/N \tag{4}$$

where S is the number of substitutions, D is the number of deletions, I is the number of insertions, C is the number of correct words and N is the number of words in the reference ($N = S + D + C$). Note that if the number of insertions is very high, the WER can be greater than 100%. Table II present the WER percentages of translation between Kachin and Rawang. Note that this WER table is calculated based on one experimental result of the 10-fold cross validation. Table II shows that HPBSMT gave the lowest WER (41.6%) for Kachin-Rawang translation and OSM gave the lowest WER (41.1%) for Rawang-Kachin translation.

From our studies, the top 10 confusion matrixes for Kachin-Rawang and Rawang-Kachin OSM machine translation can be seen in Table III and Table IV.

We also made manual error analysis on translated outputs of the best OSM model, and we found that dominant errors are different in sentence level. The followings are some common translation error patterns for PBSMT, HPBSMT and OSM:

```
### PBSMT, Kachin-Rawang ###
No. [1]
Scores: (#C #S #D #I) 7 3 2 1
REF: NÀ Í NGÀ *** LVP KÀQ GØ DĒDVM MÁ RÀ Ē
YO .
HYP: NÀ NØ NGÀ YÀ LVP KÀQ ***** NØNT NA RÀ
Ē ** .
Eval: S I D S S D

No. [2]
Scores: (#C #S #D #I) 5 2 1 0
REF: WĒ MÉ NØ BVTGWĪN TÌQ CHVNG MV:Ī .
HYP: YÀ MÉ ***** BVTGWĪN TÌQ GWĪN MV:Ī .
Eval: S D S

### HPBSMT, Kachin-Rawang ###
No. [3]
Scores: (#C #S #D #I) 2 4 2 0
REF: Àng YØP BǾĪ WĒ Í NÀ ÍÈ .
HYP: Àng **** ***** SHÀ YUP Ē LĒ .
Eval: D D S S S S

No. [4]
Scores: (#C #S #D #I) 3 5 1 0
REF: ÀNG MÀQ VNÍ GØ KARØT SHVJVNG SHĪ Ē .
HYP: ÀNG **** VNÍ SANHTAI WÀ SHAMAN VL LĒ .
Eval: D S S S S S

### OSM, Kachin-Rawang ###
No. [5]
Scores: (#C #S #D #I) 6 1 0 1
REF: SHǾNGTØNG RĪ ***** KĀDVNG TØNG ÍÈ LĒ
.
HYP: TØNG RĪ NØ KĀDVNG TØNG ÍÈ LĒ .
Eval: S I


No. [6]
Scores: (#C #S #D #I) 7 0 0 1
REF: YÀ MÉ NØ NGÀ *** SHÀ LŌ .
HYP: YÀ MÉ NØ NGÀ YÀ SHÀ LŌ .
Eval: I

### PBSMT, Rawang-Kachin ###
No. [7]
Scores: (#C #S #D #I) 2 0 0 2
REF: nyau langai *** **
HYP: nyau langai SHA GA
Eval: I I
```

TABLE I: Average BLEU and RIBES scores for PBSMT, HPBSMT and OSM

src-tgt	PBSMT	HPBSMT	OSM
kc-rw	43.071 (0.79964)	43.197 (0.80064)	**43.973 (0.80079)**
rw-kc	46.281 (0.81064)	46.129 (0.81224)	**46.597 (0.81180)**

TABLE II: Average WER% for PBSMT, HPBSMT and OSM with around 1,000 sentences (lower is better)

src-tgt	PBSMT	HPBSMT	OSM
kc-rw	42.5%	**41.6%**	42.7%
rw-kc	43.6%	42.6%	**41.1%**

TABLE III: The top 10 confusion pairs of OSM model for Kachin-Rawang machine translation

Freq	Reference ==> Hypothesis
33	wĒ ==> yÀ
26	yÀ ==> wĒ
19	lÓ ==> ÍÈ
12	chvng ==> gwĪn
10	Ē ==> ngĒ
9	ÍÈ ==> Ē
8	. ==> ?
8	kŪ ==> wĒ
7	? ==> .
7	gwĪn ==> bvtgwĪn

TABLE IV: The top 10 confusion pairs of OSM model for Rawang-Kachin machine translation

Freq	Reference ==> Hypothesis
13	u. ==> ai.
12	ai ==> le
11	. ==> ai.
11	ai. ==> .
10	ai ==> na
10	ai. ==> re.
9	na ==> ai
8	le ==> ai
7	ai. ==> rai?
7	ai. ==> sai.

No. [8]
Scores: (#C #S #D #I) 4 0 2 0
REF: dai GAW yakhkak NGA ai i?
HYP: dai *** yakhkak *** ai i?
Eval: D D

### HPBSMT, Rawang-Kachin ###

No. [9]
Scores: (#C #S #D #I) 4 1 0 1
REF: ngai gaw mi mi ** RE.
HYP: ngai gaw mi mi RE AI.
Eval: I S

No. [10]
Scores: (#C #S #D #I) 4 2 0 1
REF: dai BALL PEN langai re ** i?
HYP: dai GAW KUMGA langai re AI i?
Eval: S S I

### OSM, Rawang-Kachin ###

No. [11]
Scores: (#C #S #D #I) 6 1 0 3
REF: * dai gaw *** HKAGAWM langai re *** ai i?
HYP: N dai gaw HKA KAWK langai re NGA ai i?
Eval: I I S I

No. [12]
Scores: (#C #S #D #I) 6 1 2 1
REF: * dai GAW MAW DAW langai re nga ai i?
HYP: N dai *** *** MAWDAW langai re nga ai i?
Eval: I D D S

Where "Scores" are operation scores of the Edit Distance [27], "C" is the number of correct words, "S" is the number of substitutions, "D" is the number of deletions, "I" is the number of insertions, "REF" for reference, "HYP" for hypothesis and "Eval" is the ordered sequence of edit operations.

We found that one of the translation error patterns is paraphrasing (e.g. No. 9 and No. 5) and they are really interesting. The source (Rawang language) sentence of the example number 9 is "NGÀ NØ MĪ MĪ ÍÈ ." ("I am MĪ MĪ. " in English, "ကျွန်မ က မိမိပါ" in Myanmar language). Here, the reference is "RE" and the model output or hypothesis is "RE AI". If we only consider the meaning between reference and hypothesis, the Rawang-Kachin HPBSMT model is working well. This is because the meanings of both "RE" and "AI" are the same (i.e. ending word of a sentence, like "တယ်" in Myanmar language).

One more error type that we noticed is "word segmentation". For example, see error pattern example number 12. Here, the source (Rawang language) sentence is "WĒ MÉ NØ MODŌ TÌQ CHVNG ÍÈ ." and the meaning is "this is a car" in English. The edit distance operations are required for the word segmentation difference between reference word “MAW DAW” and the hypothesis word “MAWDAW” . In this case, the correct word segmentation is “MAWDAW” (“car” in English) and the original manual word segmentation of the reference word is wrong.

Moreover, we also found that some of the translation errors are happening at the end part of the translated sentences. For example, the normal sentence ending of Kachin language ”ai”, ”re” and question words ”rai?” and ”i?”.

X. Conclusion

This paper contributes the first PBSMT, HPBSMT and OSM machine translation evaluations from Kachin to Rawang and Rawang to Kachin. We used the 10,000 Kachin-Rawang parallel corpus that we constructed to analyze the behavior of a dialectal Kachin-Rawang machine translation. We showed that higher BLEU and RIBES scores can be achieved for Kachin-Rawang language pair even with the limited data. In fucture work, we would like to explore SMT and NMT approaches for Kachin or Jinghpaw dialect languages such as Lisu, Lashi, Rawang, Zaiwa, Lhao Vo.

Acknowledgment

We would like to thank Mr. M Jhonathan for his advice especially on manual translation between Kachin and Rawang languages. Thank you so much for the opportunity to be internship students of the Language Understanding Lab., Pyin Oo Lwin, Myanmar. Last but not least, we would like to thank Dr. May Phyo Oo (Principal, University of Computer Studies Banmaw, Myanmar) for all the help and support for this Kachin-Rawang parallel corpus building and machine translation project.

References

[1] Koehn, Philipp and Och, Franz Josef and Marcu, Daniel, “Statistical phrase-based translation,” Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1, 2003, pp. 48–54.

[2] Koehn, Philipp and Hoang, Hieu and Birch, Alexandra and Callison-Burch, Chris and Federico, Marcello and Bertoldi, Nicola and Cowan, Brooke and Shen, Wade and Moran, Christine and Zens, Richard and Dyer, Chris and Bojar, Ondřej and Constantin, Alexandra and Herbst, Evan, A. Constantin, and E. Herbst, “Moses: Open source toolkit for statistical machine translation,” Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, 2007, pp. 177–180.

[3] Koehn, Philipp, “Europarl: A parallel corpus for statistical machine translation,” Conference Proceedings: the tenth Machine Translation Summit, 2005, pp. 79–86.

[4] Ye Kyaw Thu, Andrew Finch, Win Pa Pa, and Eiichiro Sumita, “A Large-scale Study of Statistical Machine Translation Methods for Myanmar Language,” in Proceeding of SNLP2016, February 10-12, 2016.

[5] Chiang, David, “Hierarchical phrase-based translation,” Computational Linguistics 33(2), 2007, pp. 201-228.

[6] Papineni, Kishore and Roukos, Salim and Ward, Todd and Zhu, Wei-Jing, “BLEU: a Method for Automatic Evaluation of Machine Translation,” Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, ACL ’02, Philadelphia, Pennsylvania, 2002, pp. 311–318

[7] Isozaki, Hideki and Hirao, Tsutomu and Duh, Kevin and Sudoh, Katsuhito and Tsukada, Hajime, “Automatic evaluation of translation quality for distant language pairs,” Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, 2010, pp. 944-952.

[8] Win Pa Pa,Ye Kyaw Thu, Andrew Finch and Eiichiro Sumita, “A Study of Statistical Machine Translation Methods for Under Resourced Languages,” 5th Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU Workshop), 09-12 May, 2016, Yogyakarta, Indonesia, Procedia Computer Science, Volume 81, 2016,pp. 250–257.

[9] 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.

[10] Karima Meftouh, Salima Harrat, Salma Jamoussi, Mourad Abbas and Kamel Smaili, “Machine Translation Experiments on PADIC: A Parallel Arabic DIalect Corpus,” oin Proc. of the 29th Pacific Asia Conference on Language, Information and Computation, PACLIC 29, Shanghai, China, October 30 - November 1, 2015, pp. 26-34.

[11] Neubarth Friedrich, Haddow Barry, Huerta Adolfo Hernandez and Trost Harald, “A Hybrid Approach to Statistical Machine Translation Between Standard and Dialectal Varieties,” Human Language Technology, Challenges for Computer Science and Linguistics: 6th Language and Technology Conference, LTC 2013, Poznan, Poland, December 7-9, 2013, Revised Selected Papers, pp .341–353.

[12] Pierre-Edouard Honnet, Andrei Popescu-Belis, Claudiu Musat and Michael Baeriswyl, “Machine Translation of Low-Resource Spoken Dialects: Strategies for Normalizing Swiss German,” CoRR journal, volume (abs/1710.11035), 2017.

[13] John Okell, “Three Burmese Dialects,” 1981, London Oxford University press, Univeristy of London.

[14] Lucia Specia,, “Tutorial, Fundamental and New Approaches to Statistical Machine Translation,” International Conference Recent Advances in Natural Language Processing, 2011

[15] Braune, Fabienne and Gojun, Anita and Fraser, Alexander, “Long-distance reordering during search for hierarchical phrase-based SMT,” n Proc. of the 16th Annual Conference of the European Association for Machine Translation, 2012, Trento, Italy, pp. 177-184.

[16] Durrani, Nadir and Schmid, Helmut and Fraser, Alexander, “A Joint Sequence Translation Model with Integrated Reordering,” in Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, 2011, Portland, Oregon, pp. 1045-1054.

[17] Nadir Durrani, Helmut Schmid, Alexander M. Fraser, Philipp Koehn and Hinrich Schutze “The Operation Sequence Model - Combining N-Gram-Based and Phrase-Based Statistical Machine Translation,” Computational Linguistics, Volume 41, No. 2, 2015, pp. 185-214.

[18] Prachya, Boonkwan and Thepchai, Supnithi, “Technical Report for The Network-based ASEAN Language Translation Public Service Project,” Online Materials of Network-based ASEAN Languages Translation Public Service for Members, NECTEC, 2013.

[19] Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation,” Annual Meeting of the Association for Computational Linguistics (ACL), demonstration session, Prague, Czech Republic, June 2007.

[20] Och Franz Josef and Ney Hermann, “Improved Statistical Alignment Models,” in Proceedings of the 38th Annual Meeting on Association for Computational Linguistics, Hong Kong, China, 2000, pp. 440-447.

[21] Tillmann Christoph, “A Unigram Orientation Model for Statistical Machine Translation,” in Proceedings of HLT-NAACL 2004: Short Papers, Stroudsburg, PA, USA, 2004, pp. 101-104.

[22] Heafield, Kenneth, “KenLM: Faster and Smaller Language Model Queries,” in Proceedings of the Sixth Workshop on Statistical Machine Translation, WMT ’11, Edinburgh, Scotland, 2011, pp. 187-197.

[23] Chen Stanley F and Goodman Joshua, “An empirical study of smoothing techniques for language modeling,” in Proceedings of the 34th annual meeting on Association for Computational Linguistics, 1996, pp. 310-318.

[24] Och Franz J., “Minimum error rate training in statistical machine translation,” in Proceedings of the 41st Annual Meeting n Association for Computational Linguistics – Volume 1,Association for Computer Linguistics, Sapporo, Japan, July, 2003, pp.160-167.

[25] Thazin Myint Oo, Ye Kyaw Thu, Khin Mar Soe, “Statistical Machine Translation between Myanmar (Burmese) and Rakhine (Arakanese)”, In Proceedings of ICCA2018, February 22-23, 2018, Yangon, Myanmar, pp. 304-311

[26] (NIST) The National Institute of Stan- dards and Technology, Speech recognition scoring toolkit (sctk), version: 2.4.10, 2015

[27] Miller, Frederic P. and Vandome, Agnes F. and McBrewster, John, Levenshtein Distance: Information Theory, Computer Science, String (Computer Science), “String Metric, Damerau Levenshtein Distance, Spell Checker, Hamming Distance”, ISBN: 6130216904, 9786130216900, Alpha Press, 2009.