The Application of Phrase Based Statistical Machine Translation Techniques to Myanmar Grapheme to Phoneme Conversion

Ye Kyaw Thu†, Win Pa Pa‡, Andrew Finch†, Jinfu Ni§, Eiichiro Sumita†, and Chiori Hori§

† Multilingual Translation Lab., NICT, Kyoto, Japan
‡ Natural Language Processing Lab, UCSY, Yangon, Myanmar
§ Spoken Language Communication Lab., NICT, Kyoto, Japan
winpapa@ucsy.edu.mm
{yekyawthu,andrew.finch,jinfu.ni,eiichiro.sumita,chiori.hori}@nict.go.jp

Abstract. Grapheme-to-Phoneme (G2P) conversion is a necessary step for speech synthesis and speech recognition. In this paper, we attempt to apply a Statistical Machine Translation (SMT) approach for Myanmar G2P conversion. The performance of G2P conversion with SMT is measured in terms of BLEU score, syllable phoneme accuracy and processing time. The experimental results show that G2P conversion with SMT is outperformed a Conditional Random Field (CRF) approach. Moreover, the training time was considerably faster than the CRF approach.

Keywords: G2P, SMT, CRF, Phoneme, Myanmar language

1 Introduction

G2P conversion is the task of predicting the pronunciation of words given only the spelling. A grapheme is the smallest semantically distinguishing unit in a written language analogous to the phonemes of spoken languages. The correspondence between graphemes and phonemes of the Myanmar language is not as simple as one to one. The relationship between syllables and pronunciations is context dependent, depending on adjacent syllables, and there are many exceptional cases. Some syllables can be pronounced in more than 4 ways depending on the context and Part of Speech (POS) of the syllable.

This is the reason Myanmar G2P conversion cannot be performed sufficiently well using Dictionary based approaches or rule-based approaches. Some Myanmar words' pronunciation can vary across different dialects of Myanmar. We will focus only on standard Myanmar pronunciation in this paper.

We took into account Myanmar subscript words or Pali words and foreign words in this work. The transcription of foreign words pronunciation is not standardized for Myanmar language, therefore some foreign words can be written in Myanmar in more than one way and there can be an out of vocabulary (OOV) problem.

We applied phrase based SMT for sentence level Myanmar language G2P conversion. Phoneme tagged training and test sentences are prepared manually. The performance of G2P conversion on SMT was compared with CRF approach.

The rest of this paper is organized as follows. Section 2 describes Related Work for G2P and section 3 explains G2P Mapping. Preparing Training Data and Pronunciations of syllables is explained in section 4 and 5. Section 6 is Experiment of G2P with SMT and CRF, section 7 presents results of experiment, section 8 is about discussion and section 9 concludes the paper.

2 Related Work

There is only one published paper for Myanmar language G2P conversion so far. It was a dictionary-based approach and worked on only Myanmar syllables and did not consider Pali or subscript consonants [1]. The main drawback is out of vocabulary word (OOV) since it was a dictionary-based approach. G2P conversion for English and non-English languages have been proposed using rule-based, data-driven and statistical methods [8][9][10][11]. [8] compared different G2P methods and found that data-driven methods outperform rule-based methods. A novel modified Expectation-Maximization (EM)-driven G2Psequence alignment algorithm that supports joint-sequence language models, and several decoding solutions using weighted finite-state transducers (WFST) is presented in [10].

G2P conversion using SMT is proposed by [2] [3]. In [2], it shows that applying SMT gives better results than a joint sequence model-based G2P converter for French. The automatic generation of a pronunciation dictionary is proposed in [3] and it used the Moses phrase-based statistical machine translation toolkit [12] as G2P conversion.

3 Grapheme to Phoneme Mapping

The Myanmar Language Commission (MLC) Pronunciation Dictionary can be used as a basis for pronunciation mapping [4]. We found it necessary to extend the dictionary with foreign pronunciations. In the proposed mapping table there are 23 phonetic symbols for 33 consonants (some consonants share the same pronunciation, for example “ဒ”, “ဓ”, “ဍ” and “ဎ” in Table1), 87 vowels combinations and 20 special symbols for foreign word pronunciations.

Characters are grouped according to their pronunciation; the groups are un-aspirated, aspirated, voiced and nasal and are shown in Table 1. Many Myanmar syllables containing un-aspirated and aspirated consonants are pronounced as voiced consonants depending on the neighboring context.

Some foreign pronunciations have to be expressed by special vowel combinations because Myanmar pronunciations do not include some pronunciations. See Table 3. MLC dictionary was extended by defining 26 more symbols to include phoneme mappings for foreign words for example, the Myanmar phonetic representation of the foreign name “Alex” “အဲလက်(စ်)” is e:le’S (here, S is for (စ်)) and “Swift” “ဆွစ်(ဖ်)(ထ်)” is hswi’HPHT (here, HP is for (ဖ်) and HT is for (ထ်)).

Grouped consonants				
Unaspirated	**Aspirated**	**Voiced**		**Nasal**
က /k/	ခ /kh/	ဂ /g/	ဃ /g/	င /ng/
စ /s/	ဆ /hs/	ဇ /z/	ဈ /z/	ဉ/ည /nj/
ဋ /t/	ဌ /ht/	ဍ /d/	ဎ /d/	ဏ /n/
တ /t/	ထ /ht/	ဒ /d/	ဓ /d/	န /n/
ပ /p/	ဖ /hp/	ဗ /b/	ဘ /b/	မ /m/
ယ /j/	ရ /j/ or /r/	လ /l/	ဝ /w/	သ /th/
	ဟ /h/	ဠ /l/	အ /a/	

Table 1: Groups of Myanmar Consonants

4 Preparing Training Data

We built two types of data for training. The first one was a phonetic dictionary based on the MLC phonetic dictionary that contains pronunciations of 26,588 unique words. Myanmar language data from the multilingual Basic Travel Expression Corpus (BTEC) [6], which is a collection of travel-related expressions, was the second type of data used for training.

4.1 Building the Phonetic Dictionary

The phonetic dictionary was built for training the G2P conversion model by modifying entries for existing words, and by adding new words to the MLC phonetic dictionary. The following steps were applied for syllable-to-phoneme alignment to the dictionary in order:

1. Words from MLC dictionary were broken into syllables using a heuristic approach, which is 100
2. Syllables were aligned to their phonemes using a combination of rules and human annotation. Initially, single syllable words were used to align by exact match on the phoneme sequences. This was sufficient to unambiguously align about 80
3. Map MLC phonemes to the proposed phoneme set using a manually prepared conversion table.

The size of the upgraded dictionary was 28,393 unique words (2,489 unique syllables, 1,906 unique phonemes).

4.2 Sentence Selection with a Greedy Algorithm

The size of BTEC1 subset of the BTEC corpus used in these experiments was 160 K sentenced and manually phoneme-tagging all these sentences would be a time consuming task. Therefore a phonetically balanced sample was taken that contained all syllables by applying the greedy algorithm proposed in [13]. We briefly describe this algorithm below.

To select such a sentence set $\mathcal{S}$ from a large text corpus, it is necessary to define the metric of unit coverage of the sentence set. Let unit type, X, have elements $\{\mu_1^x, \mu_2^x, ..., \mu_{nx}^x\}$, where n_x is the number of elements. X can be a syllable, a diphone, or other defined unit. Assume $p(\mu_i^x)$ the occurrence frequency of μ_i^x in the text corpus. By definition, $\sum_{i=1}^{n_x} p(\mu_i^x) = 1$. The unit coverage of $\mathcal{S}$ to X, denoted by $C_{\mathcal{S}}^X$, is defined as $C_{\mathcal{S}}^X = \sum_{i=1}^{n_x} p(\mu_i^x) \times \sigma(\mu_i^x)$, where $\sigma(\mu_i^x) = 1$, if $\mu_i^x \in S$. Otherwise, $\sigma(\mu_i^x) = 0$. When given a text corpus and the size of $\mathcal{S}$ in the number of sentences, say n, the goal is to select in sentences from the text corpus to maximize $C_{\mathcal{S}}^X$. In this paper, four types of units are considered, namely, syllable, di-phone spanning two syllables, tri-phone.

The main steps of the algorithm are briefly described as follows.

Step 1 : Calculate the occurrence frequency of units in the text corpus and set none to $\mathcal{S}$.

Step 2 : Scan the whole text corpus to select one sentence and add the sentence to the current sentence set $\mathcal{S}$. The selected sentence is that which maximizes the unit coverage of $\mathcal{S}$ according to the following priority:

(1) Maximizing $C_{\mathcal{S}}^{syllable}$ (or simply denoted by $C_{\mathcal{S}}^{syl}$).
(2) Maximising $C_{\mathcal{S}}^{diphone\ spanning\ two\ syllables}$, if (1) is satisfied.
(3) Maximizing $C_{\mathcal{S}}^{diphone}$, if (1)-(2) are satisfied.
(4) Maximizing $C_{\mathcal{S}}^{triphone}$, if (1)-(3) are satisfied.

In this way, the algorithm can find the best sentence set that simultaneously maximizes the coverage of syllables, di-phone spanning two syllables, di-phones, and tri-phones in the priority mentioned above.

Step 3 : Halt, if a predefined set size is reached. Otherwise, repeat Step 2.

We extracted 5,276 sentences from BTEC1 and used them for training using SMT and CRF models. The 5276 sentences were tagged with their phonemes manually. The selected sentences set contain foreign names and which should allow for the coverage of non-Myanmar words.

The pronunciation is formed from syllables and syllable boundaries have to be defined since Myanmar language is written continuously. Words from MLC Dictionary and selected sentences were first broken into syllables using a heuristic approach [7]. Then each syllable was labeled with its phoneme based on MLC Dictionary. Labeling phonemes on selected 5,276 sentences was manually done by three Myanmar native speakers. Some phoneme tagged sentences are shown in Figure 1 and it can be seen that pronunciations of some same syllables are different.

ကျေး ဇူး တင် ပါ တယ် ။
kyei: zu: tin ba de pm
(Thank you.)

မင်း စဉ်း စဉ်း စား စား လုပ် တတ် တယ် ။
min: sin: zin: sa: za: lou' ta' te pm
(You usually do carefully.)

မိ ချောင်း သား ရေ သား က ဘာ လဲ ။
mi. gyaun: tha- jei dha: ga. ba le pm
(what is crocodile skin?)

Fig. 1. Phoneme tagged sentences

5 Pronunciations of Syllables

Pronunciations of Myanmar syllables can be different from the original pronunciation of orthographic structure. The following two sub-sections explain the original pronunciation of syllables and how they can change according to their context.

5.1 Contextually Independent Pronunciation

This section explains how the pronunciation of Myanmar syllables is normally derived from orthographic structure. Myanmar syllables are generally composed of consonants and (zero or more) vowel combinations starting with a consonant. Here, vowel combinations can be single vowel, sequences of vowels starting with a consonant that modifies the pronunciation of the first vowel. The pronunciations of consonants when they are combined with vowels are shown in Table 2.

အီ i	အိ i.	အီး i:	အစ် i'	အင် in	အင့် in.	အင်း in:
အေ ei	အေ့ ei.	အေး ei:	အိတ် ei'	အိန် ein	အိန့် ein.	အိန်း ein:
အယ် e	အယ့် e.	အဲ e:	အိုက် ai'	အိုင် ain	အိုင့် ain.	အိုင်း ain:
အာ a	အာ့ a.	အား a:	အတ် a'	အန် an	အန့် an.	အန်း an:
အော် o	အော့ o.	အော o:	အောက် au'	အောင် aun	အောင့် aun.	အောင်း aun:
အူ u	အု u.	အူး u:	အွတ် u'	အွန် un	အွန့် un.	အွန်း un:
အို ou	အို့ ou.	အိုး ou:	အုပ် ou'	အုန် oun	အုန့် oun.	အုန်း oun:

Table 2: Examples of vowel combinations and their pronunciations

In general, the pronunciation of syllables can be obtained directly from the pronunciation of these components. All of the pronunciations of consonants are shown in Table 1 and some example pronunciations of vowel combinations are shown in Table 2. The pronunciation of full syllable is a concatenation of the pronunciations of each component. The pronunciations do not modify each other. This type of pronunciation will be referred as the “standard pronunciation”. Table 3 shows examples of standard pronunciations of some syllables according to their composition of consonant and vowels.

Syllable	Consonant+Vowel	Standard Pronunciation
သ	သ	tha.
တင်း	တ+င်း	t+in:
စာ	စ+◌ာ	s+a
ပိုင်း	ပ+◌ိုင်း	p+ain:
ခြေ	ခြ+ေ◌	ch+ei
ပု	ပ+◌ု	p+u
လဲ	လ+◌ဲ	l+e:
မင်	မ+င်	m+in
ကြောင်	ကြ+ေ◌ာင်	ky+aun

Table 3: Standard pronunciation of syllables

5.2 Contextually Dependent Pronunciations

Some Myanmar syllables do not conform to these standard rules of pronunciation. The pronunciation of the syllables can depend on the context of syllables.

No.	Words	Standard	Correct
1	သတင်းစာ (newspaper)	tha. tin: sa	dha- din: za
2	ပိုင်းခြေ (denominator)	pain: chei	pain: gyei
3	ပုလဲ (pearl)	pu. le:	pa- le:
4	ပညာ (knowledge)	pa. nja	pjin nja
5	မင်ကြောင် (tatoo)	min kyaun	mhin gyaun

Table 4: Examples pronunciations of some words

Differences between standard pronunciations and correct pronunciations of some words are shown in Table 4 as examples. It can be also seen in Table 2 that pronunciations of some same syllables are different depend on the context.

In [5], 10 patterns are proposed to capture the dependencies. Most of the patterns changed unaspirated or aspirated syllables to their voiced form. The first word in Table 4 is one pattern of deviation to standard pattern. The pattern is of that word’s pronunciation is “A syllable’s pronunciation can be changed if the syllable before it was changed”. This phenomenon can cause a cascade of changes that can affect several syllables. The next pattern is changing the pronunciation of successive aspirated or aspirated syllable to voiced sound if the vowel combination of first syllable is (/in/, /an/, /e/, /aun/, /ein/, /un/). The second word in Table 4 is an example of that pattern. The pronunciation of some vowel combinations is occasionally non-standard. An example of that pattern is the third word in Table 4. Vowel sounds of some syllables are omitted and sometime nasal vowel is added to the pronunciation. The fourth and fifth words in Table 4 follow that pattern. Another pattern relates to compound words. If a noun and verb combine to form a noun phrase, the final syllable (unaspirated or aspirated) is changed to the voiced form.

We found every pattern has exceptions and the amount of exceptions is not small. In one pattern, the number of exceptions is 2,659 in 6,446 occurrences of that pattern and number exception of another pattern is 344 of 4,817 of occurrences. There are 224 exceptions in out of 388 occurrences in the pattern of omission of vowel sound. From this fact, it is obvious that only rule based or dictionary based approaches cannot predict correctly for Myanmar G2P conversion.

6 Experiments

6.1 Data Settings

We use three training data settings: extended version of MLC Dictionary (28,393 unique words) that we prepared (see Section 4.1), selected 5,276 sentences (see Section 4.2) with a greedy algorithm and combination of two of them. Training data were mapped to our phoneme symbols to create the training data for building CRF models and phrase based translation models.

We prepared 4 open test sets that were randomly selected from BTEC corpus; 3 sentence level test sets (Test set 1, Test set 2, Test set 3) and 1 word level test set (Test set 4). Each sentence level test set has 500 sentences. In detail, the unique number of Test sets are: Test set 1 contains 831 syllables and 837 phonemes, Test set 2 contains 836 syllables and 844 phonemes, and Test set 3 contains 824 syllables and 822 phonemes. Word level test set 4 contains 414 words, 461 syllables and 480 phonemes.

6.2 Conditional Random Field (CRF) Models

We used the CRFsuite tool [15] for training and testing CRF models. The feature set consisted of unigrams and bigrams of syllables, and unigrams, bigrams and trigrams of pronunciation change labels for each feature, and is shown below:

s[t-2], s[t-1], s[t], s[t+1], s[t+2]
s[t-1]|s[t], s[t]|s[t+1]
l[t-2], l[t-1], l[t], l[t+1], l[t+2]
l[t-2]|l[t-1], l[t-1]|l[t], l[t]|l[t+1], l[t+1]|l[t+2]
l[t-2]|l[t-1]|l[t], l[t-1]|l[t]|l[t+1], l[t]|l[t+1]|l[t+2]

Where s[t] is the syllable at position t (t being the position of the syllable being labeled), and l[t] is the label at position t; and s[t-1]s[t] is a bigram of syllables, and so on. In addition the model includes transition features for up to bigrams of phonemes.

6.3 Phrase Based Translation Models

We use the phrase based SMT system in Moses for training machine translation model [12]. The Myanmar source segmented by syllable segmentation method is aligned to the syllable level phoneme using GIZA++ [16]. The alignment is symmetrized by grow-diag-final-and heuristics [17]. The lexicalized reordering model is trained with the msd-bidirectional-fe option [18]. We use SRILM to training 5-gram language model with interpolated modified Kneser-Ney discounting on phoneme training data [19] [20]. In decoding, we adopt the default settings of the Moses decoder. Since the size of manually phoneme tagged data is small, tuning was not done for all SMT experiments in this paper.

6.4 Evaluation Criteria

We used two criteria for evaluation; Bilingual Evaluation Understudy (BLEU) [14] for SMT and phoneme accuracy for comparison between CRF and SMT approaches.

7 Results

The results of the SMT experiment with dictionary model, selected sentences model and combination of dictionary and selected sentences model are shown in Table 5. From the results, generally, translation model trained with sentence level give better G2P translation than word level for Test Set 1, 2 and 3. The highest BLEU score 86.29 achieved from the Dictionary+Sentence SMT model.

Test Data	**Dictionary**	**Sentence**	**Dict+Sentence**
Test Set 1	37.15	**74.63**	74.59
Test Set 2	34.70	73.73	**74.18**
Test Set 3	38.55	75.33	**75.57**
Test Set 4	80.34	79.17	**86.29**

Table 5: Test set BLEU score of Myanmar G2P

The comparison in terms of phoneme accuracy between CRF and SMT approaches is shown in Table 6. It can be seen clearly that G2P conversion using SMT approaches outperformed that of CRF.

We also measured training time difference between CRF and SMT approaches. Here, we used three different servers with similar specification. Training time for

	CRF			SMT		
Test-Data	**dictionary**	**sentence**	**dict+sent**	**dictionary**	**sentence**	**dict+sent**
Test Set 1	50.48	73.56	74.21	65.34	**89.66**	89.59
Test Set 2	49.60	73.82	74.36	63.64	89.24	**89.45**
Test Set 3	51.31	74.55	75.17	65.69	89.94	**90.12**
Test Set 4	75.93	72.71	77.71	92.79	91.85	**94.29**

Table 6: Phoneme accuracy of CRF and SMT approaches

dictionary model, sentence model and dictionary + sentence model can be seen in Figure 2. From this results, SMT approach considerably faster than CRF and able to train in less than one minute.

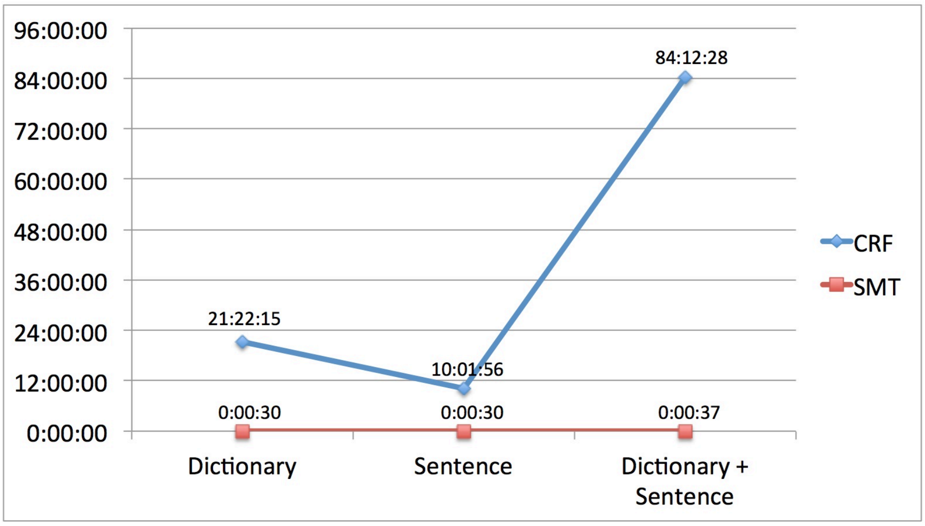

Fig. 2. Training time comparison between CRF and SMT approaches

8 Discussion

There are 1816 unique syllables and 1829 unique phonemes in BTEC and the coverage of syllables and phonemes of selected 5276 sentences from BTEC is shown in Table 7.

	Syllables	**Di-phone**	**Tri-phone**
Unit	1,833	5,121	35,295
Coverage	99.77%	90.95%	88.78%

Table 7: Syllable and phone coverage for the selected 5,276 sentences

There is usually more than one pronunciation for some syllables and all pronunciations of each syllable are included in dictionary. From the results presented in Section 7, it can be concluded that models training with a dictionary is not enough for sentence level G2P for Myanmar language. Since, pronunciations of syllables can be contextually dependent such as some particles and conjunctions. Figure 3 shows an example sentence where SMT can predict the right pronunciation of the respective syllable.

Myanmar Sentence: ဒါအတော်ပဲ။
Syllable Breaked: ဒါ အ တော် ပဲ ။
Dictionary Model Output: da a- (do)(pe:) pm
Sentence Model Output: da a- to be: pm
Dictionary+Sentence Model Output: da a- to be: pm

Fig. 3. An example of three SMT model outputs

In predicting new compound words (OOV), all CRF models (dictionary, selected sentences model, dictionary+selected sentences model) predicted as its standard pronunciation but the sentence model and dictionary+sentence SMT model can predict the correct phonemes. An example OOV word, အ ခက် ခဲ ဆုံး (the most difficult)

standard pronunciation: a- khe' khe: **hsoun:**
correct pronunciation: a- khe' khe: **zoun:**

where "ခက် ခဲ" is the stem word in the dictionary it combines with an affix "အ" and suffix "ဆုံး" to form an adjective. Some OOV words that are predicted erroneously by all CRF models and cannot be predicted by the SMT models are shown below with their correct pronunciations.

တစ် လုံး လောက် (one piece please) ta- loun: lau'
တစ် နေ ရာ ရာ (somewhere) ta- nei ja ja
တို အောင် (to be shorter) tou aun

ကမ် ဘော ဒီး ယား (Cambodia) kan bo: di: ja:
နှ ယူးစ် ဝိခ် (Newsweek) na- ju:S wi.KH

But a foreign name ရုရှား (Russia) pronounce as ra- sha: that appeared only once in the training corpus is predicted correctly by the sentence SMT model but not by the CRFs and other SMT models.

An example imperative sentence that ends with stem verb without suffix and its outputs from different models are shown below:

အောက် ကို ကြည့် ။
Look below.
au' kou kyi.

CRF (dictionary model): au' kou **za- ga**
CRF (sentence/dictionary+sentence model): au' kou **ba**
SMT (dictionary model): au' kou **ကြည့်**
SMT (sentence/dictionary+sentence model): au' kou **kyi.**

The pronunciation of ကြည့် was predicted correctly by the SMT sentence and dictionary+sentence models. In the sentence level test sets 1, 2 and 3, the pronunciation of some syllables that have two phonemes, for example

for example

ကျွန်ုပ်(I)=> kya- nou'

is predicted correctly by all SMT models but not by the CRF models.

The conjunctions and some (unaspirated/aspirated to voiced) changed pronunciations are correctly predicted by the sentence CRF models and sentence SMT models.

The outputs of the CRF models and SMT models for an input Myanmar sentence along with its meaning and correct pronunciation are shown in below.

ဒုက် ခ ပဲ ပြော တဲ့ စ ကား က မ တူ ဘူး ။
So bad, we are speaking in different languages.
dou' kha. be: pjo: de. za- ga: ga. ma- tu bu:

CRF (dictionary model):
dou' kha. pe: pjo: te. za- ga: ka- ma- tu bu:

CRF (sentence/dictionary+sentence model):
dou' kha. be: pjo: de. za- ga: ga. ma- tu bu:

SMT (dictionary model):

dou’ kha. pe: pjo: te. za- ga: ka- ma- du bu:

SMT (sentence/dictionary+sentence model):
dou’ kha. be: pjo: de. za- ga: ga. ma- tu bu:

In the above example, all models except dictionary models can predict the correct pronunciations.

In all the CRF models, the dictionary+sentence model has achieved the highest accuracy but among the SMT models, the sentence model achieved similar accuracy to the dictionary+sentence model. This indicates that the sentence SMT model can work well without a dictionary since the training sentences are selected to cover all syllables and phonemes from the original BTEC corpus. One major advantage of using SMT rather than a CRF model is speed. The SMT model proved to be considerably faster in both decoding and training, making the approach far more practicable.

9 Conclusion

In this paper, we presented G2P conversion results applying phrase based SMT. The highest BLEU score 86.29 was achieved from training only with a dictionary plus selected 5,276 sentences. All the results using SMT outperformed CRF approaches in terms of phoneme accuracy. Furthermore, our experiments have shown that the SMT approach also has a great advantage at the sentence level. In future work we hope to extend our SMT experiments with extended phoneme tagged data and also with other syllable based languages such as Thai, Khmer.

Acknowledgment

We thank Ms. Aye Mya Hlaing (UCSY, Yangon, Myanmar) and Ms. Hay Mar Soe Naing (UCSY, Yangon, Myanmar) for their help in phoneme tagging and checking for MLC dictionary and selected 5,276 sentences.

References

1. Ei Phyu Phyu Soe, “Grapheme-to-Phoneme Conversion for Myanmar Language”, 11th International Conference on Computer Applications (ICCA), pp. 195–200, Myanmar (2013)
2. Antoine Laurent, Paul Deleglise, and Sylvian Meignier, “Grapheme to phoneme conversion using an SMT system”, INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association, pp. 708–711, Brighton, United Kingdom, (September 6-10, 2009)
3. Panagiota Karanasou and Lori Lamel, “Automatic Generation of a Pronunciation Dictionary with Rich Variation Coverage Using SMT Methods”, In Proceedings of CICLing, Part II, LNCS 6609, pp.506–517, (2011)
4. Myanmar-English Dictionary, Department of the Myanmar Language Commission, Yangon, Ministry of Education, Myanmar (1993)

5. Myanmar Thadda, Myanmar Language Commission, Ministry of Education, Myanmar (2005)
6. Genichiro Kikui, Seiichi Yamamoto, Toshiyuki Takezawa, and Eiichiro Sumita, "Comparative Study on Corpora for Speech Translation", IEEE Transactions on Audio, Speech and Language, 14(5), pp.1674–1682, (2006)
7. Jinfu Ni, Toshio Hirai, and Hisashi Kawai, "Constructing A Phonetic-Rich Speech Corpus while Controlling Time- Dependent V oice Quality V ariability for English Speech Synthesis", In Proceedings of ICASSP, Volume:1, pp. I- 881–I-884 (2006)
8. R. I. Damper, Y. Marchand, M.J. Adamson and K. Gustafson, "A Comparison of Letter-to-Sound Conversion Techniques for English Text-to-Speech Synthesis ", In Proceeding of the Institute of Acoustics 20(6), pp. 245- 254 (1999)
9. A. W. Black, Kevin Lenzo and Vincent Pagel, "Issues in building general letter to sound rules", 3rd ESCA on Speech Synthesis (1998)
10. J. R. Novak, Nobuaki Minematsu, Keikichi Hirose, "WFST-based Grapheme-to-Phoneme Conversion: Open Source Tools for Alignment, Model-Building and Decoding", In Proceedings of the 10th International Workshop on Finite State Methods and Natural Language Processing, Donostia–San Sebastían, pp. 45–49, (2012)
11. Standley F. Chen, "Conditional and joint models for grapheme-to-phoneme conversion", Eurospeech (2003)
12. P. Koehn, et.al., "Moses: Open source toolkit for statistical machine translation", In Proceedings of the ACL, pp. 177–180, (2007)
13. Ye Kyaw Thu, A. Finch, Y. Sagisaka, and E. Sumita, "A study of Myanmar word segmentation schemes for statistical machine translation," In Proceedings of ICCA, Myanmar, pp. 167–179 (2013)
14. Kishore Papineni, Salim Roukos, Todd Ward, and Wei Jing Zhu, BLEU: a Method for Automatic Evaluation of Machine Translation. In Proceedings of ACL2002, pp. 311–318, Philadelphia, USA, (2002)
15. Naoki Okazaki, "CRFsuite: a fast implementation of Conditional Random Fields (CRFs)", http://www.chokkan.org/software/crfsuite/, (2007)
16. Franz Och and Hermann Ney, "Improved Statistical Alignment Models", In Proceedings of 38th Annual Meeting on Association for Computational Linguistics, pp. 440–447, (2000)
17. P. Koehn, F. J. Och, and D. Marcu, "Statistical phrase- based translation", In Proceedings of HTL-NAACL, pp. 48–54, (2003)
18. C. Tillmann, "A Unigram Orientation model for Statistical Machine Translation", In Proceedings of HTL-NAACL, pp. 101–104, (2004)
19. A. Stolcke, "SRILM-An Extensible Language Modeling Toolkit", In Proceedings of ICSLP, pp. 901–904, (2002)
20. S. F. Chen and J. Goodman, "An Empirical Study of Smoothing Techniques for Language Modeling", In Proceedings of the 34th annual meeting on Association for Computational Linguistics, pp. 310–318, (1996)