

Myanmar-English Bidirectional Machine Translation System with Numerical Particles Identification

Yin Yin Win

University of Computer Studies, Mandalay, Myanmar
E-mail: yinyinwin.mdy@gmail.com

Aye Thida

University of Computer Studies, Mandalay, Myanmar
E-mail: ayethida.royal@gmail.com

Abstract—This paper the development of Myanmar-English bidirectional machine translation system is implemented applying Rule based machine translation approach. Stanford and ML2KR parsers are used for preprocessing step. From this step, parsers generate corresponding parse tree structures. Used parsers generate corresponding CFG rules which are collected and created as synchronous context free grammar SCFG rules. Myanmar language can be written free order style, but it must be verb final structure. Therefore, CFG rules are required for reordering the structure of the two languages. After that tree to tree transformation is carried on the source tree structure which corresponds with used parser (Stanford parser or ML2KR's parser). When source parse tree is transformed as target parse tree, it is changed according to the SCFG rules. And then system carries out the morphological synthesis. In this stage, we need to solve only for English to Myanmar machine translation because Myanmar language is morphologically rich language. Therefore, particles for Myanmar language can be solved in this system by proposed algorithm. After finishing morphological synthesis, this system generates meaningful and appropriate smoothing sentences.

Index Terms—RBMT, SCFG, Tree to Tree Transformation.

I. INTRODUCTION

Machine translation system is one of Natural Language Processing (NLP) tasks which are modern computational technologies.

The first part of the machine translation is syllable and word segmentation. Space delimited languages are easy to tokenize or segment words in the sentence. However, some language such Japanese, Chinese, Thai, India and Myanmar are not clearly delimited by a space. Thus, syllable segmentation is essentially needed as text preprocessing step.

The second part of the machine translation is part of speech (POS) Tagging. POS tagging is the process to

label (tag) each word with its correct part of speech. When the POS tagging process is performed, corpus or lexicon is needed to obtain correct POS tag.

Moreover, Reordering is important reverse words from translation process for Myanmar-English machine translation to get proper target sentence. The translation will not preserve as the same word order between Myanmar and English because of their grammatical structure. Reordering is carried out with the help of parse trees. The parser could be dependency parser or Context Free Grammar (CFG) based parser. The reordering rules are written based on the parser used.

This system applies the existing processes of NLP resources. Stanford Parser [9] is used for preprocessing of the English sentence. In Stanford Parser, it accepts English sentence as input and generates parse tree as output. ML2KR [10] is used for preprocessing of the Myanmar sentence. ML2KR accepts Myanmar sentence as input and generates parse tree as output.

When source language translates to target language, the problem is different structure of both languages. The major work of this research is changing source parse tree to target parse tree according to the target language's grammar rules. Therefore reordering is needed to obtain correct grammar structure of the target language. Context Free Grammar (CFG) rules of source and target languages are needed to generate Synchronous Context Free Grammar (SCFG) rules to be easy to change the structure.

When the translation process is implemented, bilingual dictionary or bilingual corpus is also mainly needed to generate target meaning. When the source sentence is translated to the target sentence, it is needed to consider morphological synthesis because Myanmar language is morphologically rich language. After obtaining the target words meanings and synthesizing morphologically, system generates target sentence from the target parse tree.

The remaining parts of the paper are organized as follows: the works concerning machine translation system are presented in section 2, theory background of machine translation is discussed in section 3, and

architecture of the proposed system of bidirectional translation system is described in section 4. In addition, how to parse a input sentence, how to change tree structure, how to specify the target words, how to synthesize morpheme and how to generate target sentence are described in this section 4. Evaluation results are expressed in section 5 and section 6 concludes the paper.

II. RELATED WORK

Language is very important part of the communication. There are many different languages spoken in this world among which English is the global language. The most of the information is available in English. Mr.Uday C. Patkar et.al (2012) introduced mechanism which converts multi sentences, question sentences of English to Sanskrit text to speech conversion. They stated that the model consists of array of translation rules to translate from source to target sentence, which is the frame of Rule based Machine Translation System [4].

Fai Wong et.al (2011) described the application of MT based on Constraint Synchronous Grammar (CSG) in devices with limited re-sources. This paper presented the application of Constraint Synchronous Grammar (CSG) formalism to MT for handheld devices [5].

Shibli Syeed Ashrafi et.al (2013) proposed a bi-lingual MT system for Bangla translation of an English simple assertive sentence employing structural analysis using grammatical rule-based approach in the form of context-free grammars (CFGs) [6].

Khaled Shaalan et.al (2010) described the development of a novel English-Arabic bi-directional rule-based transfer MT tool in the agriculture domain [3].

R. Harsha Wardhan (2011) developed a rule-based machine translation system for English to Malayalam language pair. Their system took in the English sentence as input and parse with the help of Stanford Parser [11].

T. T. Zin et.al (2011) presented Myanmar phrases translation model with morphological analysis. The system was based on statistical approach. In statistical machine translation, large amount of information was needed to guide the translation process. When small amount of training data was available, morphological analysis was needed especially for morphology rich language. Bayes rule was also used to reformulate the translation probability of phrase pairs. Experiment results showed that proposed system could improve translation quality by applying morphological analysis on Myanmar language [7].

III. MACHINE TRANSLATION TECHNIQUES

Machine Translation refers to the use of computers to automate some of the tasks or the entire task of translating between human languages. The major machine translation techniques are:

- 1) Statistical Machine Translation (SMT),
- 2) Example Based Machine Translation (EBMT) and

- 3) Rule Based Machine Translation (RBMT).

A. Statistical Machine Translation

The statistical machine translation (SMT) is a machine translation paradigm where translations are generated on the basis of statistical models whose parameters are derived from the analysis of bilingual text corpora [8][12]. The SMT is a corpus based approach, where a massive parallel corpus is required for training the SMT systems. The SMT systems are built based on two probabilistic models: language model and translation model. The advantage of SMT system is that linguistic knowledge is not required for building them. The difficulty in SMT system is creating massive parallel corpus [4][13].

B. Example Based Machine Translation

The example based machine translation (EBMT) is the corpus based approach without any statistical models. The example based systems are trained with the parallel corpus of example sentences, similar to SMT systems. The example based systems generally don't learn from the corpus. They store the parallel corpus and uses matching algorithms to search and retrieve the sentences. The translation memories (TM) are built to aid the human translators by serving as an assisting tool for translation. The advantage of translation memories are easy to implement and linguistic knowledge is not required [4][13].

C. Rule Based Machine Translation

The rule based machine translation system translates the source text into target text by a set of linguistic rules. Three techniques of machine translation – Direct, Interlingua and Transfer based are applicable to rule based machine translation system. The rule based machine translation system is developed by hand coded rules for translation. The system requires good linguistic knowledge to write the rules and a bilingual dictionary is also needed. Other MT systems like SMT and EBMT requires huge parallel corpus for training. The rule based systems are highly suited for translation of Myanmar-English Languages because the bilingual dictionary could be collected easily compared to parallel corpus and the rules could also be written well with the help of linguists. The rule based system which has been developed follows the transfer based approach of reordering rules. The drawback of rule based system is that the system is confined with the rules and the rules will evolve with the language over time [4][13].

- 1) *Direct Approach*: Direct approach involves in four stages to translate any language to other language. Morphological analysis can be done i.e., identified the tense for the verb then Identify the constituents and Reorder the constituents based on target. Replace the source words to target with the help of dictionary. But direct approach in not a minimal structure and semantic analysis also won't produce a long term solution for MT [2][14].

- 2) *Transfer Approach*: Transfer model involves three stages: analysis, transfer and generate. Analyze the source sentence, transfer the structure of source sentence to the structure of target sentence finally translate the word, number, gender in the target words. But in this approach n generating components, n analysis components and n(n-1) transfer components are needed for n language translation, it will increase memory and working principle [2][15].

English structural order is SVO (Subject-Verb-Object) whereas Myanmar structural order is SOV. So the system needs to use rules and then parse tree yields the structure of a sentence. On the basis of the structural differences between the source and target language, a transfer based approach is applied by tree to tree transformation.

- 3) *Interlingua Approach*: The Interlingua approach considers MT as a two stage process: Extracting the meaning of a source language sentence in a language-independent form, and, Generating a target language sentence from the meaning. In this approach, burden on the analysis and generation components increases. Have to choose between various possible parses for a sentence, identify the universal concepts that the sentence refers to, and understand the relations between various concepts expressed in the sentence [2].

IV. SYSTEM ARCHITECTURE DESIGN

A Rule based machine translation system for Myanmar-English language pair has been developed. System architecture of this system is shown in Fig. 1. The system accepts Myanmar or English sentence as an input. By using Stanford or ML2KR's parser, input sentence is parsed to generate parse tree structure with respective POS Tags. Then the system performs tree to tree transformation by using SCFG rules. Bilingual Lexicon is used to specify source words to target words. In the morphological synthesis, particles or classifiers are mainly considered for articles or cardinal numbers because Myanmar language is morphologically rich. Finally, Target sentence is generated.

A. Parsing Input Sentence

Parser plays as an important role in machine translation system. In this system, Stanford parser and ML2KR's parser are used for preprocessing step of an input source sentence. Stanford parser is for parsing input English sentence. ML2KR's parser is for parsing the given input Myanmar sentence. A parse tree structure is generated with respect to the parser. Moreover, CFG rules are gotten by using parsers. Thus, SCFG rules are generated for tree to tree transformation via CFG rules. Example original parse trees of these parsers are shown in Fig. 2 and Fig. 4.

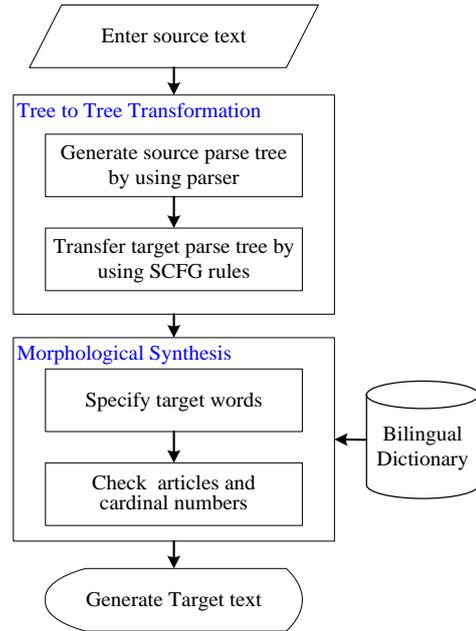


Fig.1. System Architecture.

- 1) *Context-Free Grammar*: There are four categories in Context Free Grammar. They are
 - **Set** of non terminal symbols- grammatical categories
 - **Set** of terminal symbols- words
 - **Set** of Productions- (unordered) (rewriting) rules and
 - **Distinguished** symbol- start symbol.

According to the Stanford parser, the CFG rules of an English sentence are shown in the following as an example which is “She is sleeping”.

- ROOT → S
- S → NP + VP
- NP → PRP
- PRP → She
- VP → VBZ + VP
- VBZ → is
- VP → VBP
- VBP → sleeping

According to the ML2KR's parser, the CFG rules of a Myanmar sentence are shown in the following as an example which is “သူမသည် အိပ်စက်နေသည်”.

- S → NP_NOM+VEND
- NP_NOM → NP+PREP_NOM
- NP → PRON_PERSON
- PRON_PERSON → သူမ
- PREP_NOM → သည်
- VEND → V
- V → အိပ်စက်နေသည်

2) *Synchronous Context-Free Grammar for Tree Transformation*: Synchronous Context Free Grammars (SCFG) is the mapping between two context free rules that specify how source language structures correspond to target language structures. It is a kind of context free grammar that generates pair of strings: English CFG rules and Myanmar CFG rules pairs. These rules are used to adjust the structures of the two languages. Example SCFG rules are shown in Table 1 and Table 2.

Table 1. Example SCFG rules for English Parse Tree

Phrases	E-CFG	M-CFG
NP	NP PP	PP NP
NP	DT NNS	NNS
NP	NNP	NNP
NP	PRP\$ NN	PRP\$ NN
VP	VBZ VP	VP
VP	VB NP PP	PP VB
VP	VBG PP	PP VBG
VP	VBZ NP	NP VBZ
VP	VBD PRT NP	NP PRT VBD
VP	VBP NP PRT PP	PP VBP
PP	IN NP	NP IN
PP	TO NP	NP TO

Table 2. Example SCFG Rules for Myanmar Parse Tree

Phrases	M-CFG	E-CFG
S	NP_NOM NP_ARRIVAL VEND	NP_NOM VEND NP_ARRIVAL
S	NP_NOM NP_COMPLEMENT VERB_STATIVE	NP_NOM VERB_STATIVE NP_COMPLEMENT
NP_NOM	NP PREP_NOM	NP
NP_OBJ	NP PREP_OBJ	NP
NP_ARRIVAL	NOUN_ARTIFACT PREP_ARRIVAL	PREP_ARRIVAL NOUN_ARTIFACT
NP_DEPARTURE	NOUN_ARTIFACT PREP_DEPARTURE	PREP_DEPARTURE NOUN_ARTIFACT

B. Tree to Tree Transformation

Tree to tree transformation is carried out with the help of parse trees. The parser could be dependency parser or Context Free Grammar (CFG) based parser. The reordering rules or SCFG rules are generated based on the parser used in machine translation system.

Syntactic structure of the source English text changes to Syntactic structure of the target Myanmar text. The syntactic information of source sentence from Stanford parser is produced as a source parse tree. The syntactic pattern of source sentence maps with source rule, and then the corresponding target rule is changed. The source parse tree is modified with respect to the target rule as a target parse tree.

An English sentence is composed of NP (Noun phrase) and VP (Verb Phrase) by Stanford Parser. Noun phrase is also composed of NP only, NP and PP (Preposition Phrase) or NN and PP, and so on. Moreover, Verb phrase contains different phrases types: VP, NP, ADJP, ADVP, and so on. The phrases are composed of different types. Therefore, they are recursively changed by using SCFG rules.

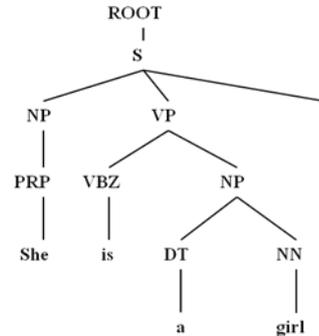


Fig.2. Parse Tree of English Sentence

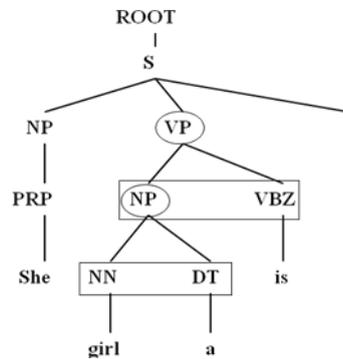


Fig.3. Parse Tree of English Sentence after Recursively Transforming

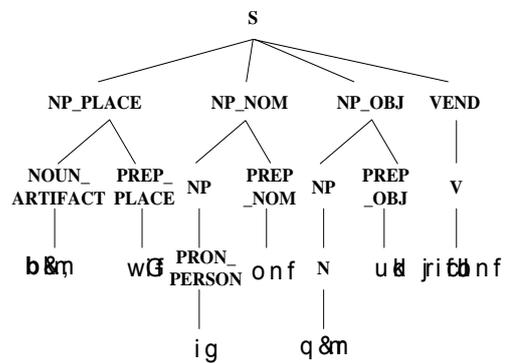


Fig.4. Parse Tree Structure of Myanmar Sentence

However English sentence structure is SVO (Subject-Verb-Object). Myanmar sentence structure is different from English Sentence Structure. Myanmar sentence structure is SOV (Subject-Object-Verb) but the structure may be OSV (Object-Subject-Verb). So Myanmar sentence structure can be written free order but its structure is verb final structure. So, phrases in a sentence are needed to reorder as phrase level transformation. Moreover, Myanmar phrase structure is not the same with

English phrase structure. So, words are needed to change as word level transformation. After transforming tree structure, target tree structures are shown in Fig. 3 and Fig. 5.

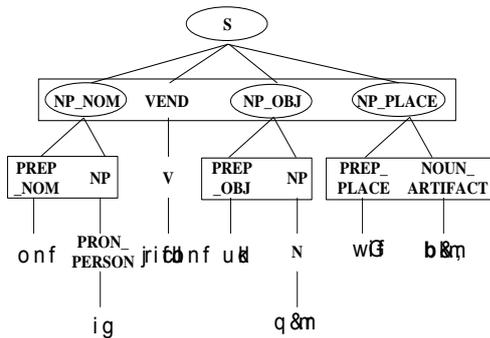


Fig.5. Parse Tree Structure after Phrase Level and Word Level Transforming

C. Target Word Specification from Bilingual Lexicon

The lexicon contains words and their corresponding meanings. The bilingual lexicon has the words in one language and their meanings in the other. English-Myanmar bilingual lexicon is used in the machine translation system for translating the English words to equivalent Myanmar words and Myanmar words to equivalent English words.

English-Myanmar bilingual lexicon is used to specify the target words. At that time, we need to justify part of speech of English words with fields of this lexicon. There are nine parts of speech in Myanmar Grammar. In this lexicon, there are two fields for noun words. They are plural noun and singular noun. ‘Engword_S’ field is for singular noun. ‘English_P’ field is for plural noun. Verb is classified four fields: simple present tense, simple past tense, present participle tense and past participle tense. ‘EngPresentSimpleS’ field is for singular verb. ‘EngPresentSimpleP’ field is for plural verb and base. ‘EngPastSimple’ field is for simple past tense. ‘EngPresentParticiple’ field is for present participle tense. ‘EngPastParticiple’ field is for past participle tense. ‘Engword’ field is for other part of speech: adverb, adjective, conjunction, particle, pronoun, preposition and interjection. ‘MMword’ field is for Myanmar words. In addition, seven groups for nouns are defined in this lexicon. Defining these group helps in solving ambiguities of preposition and numerical classifiers. These groups are Person, Animal, Thing, Instrument, Vehicle, Time, and Place.

D. Morphological Synthesis

In order to improve smooth translation, morphological synthesis is essentially needed because Myanmar language is morphologically rich. Generally, subject in Myanmar sentence is followed by subject particle such as “သည်” or “မှာ”. Object in Myanmar sentence is followed by object particle “ကို”. Moreover, plural noun is followed by particle “များ” and adjective is followed by “သော”. Verb markers are “သည်” or “ပါ”.

In this stage, the proposed system mainly considers for numerical classifiers or particles of the noun phrases. When English words are translated for articles (a, an) and cardinal numbers, we translate Myanmar words plus clue words (ခု, တောင်, ယောင်, etc). Thus, an algorithm for numerical classifier identification is also proposed in this system as shown in Fig. 6 and Fig. 7. The proposed classifier identification algorithm supports for 64 classifiers that are commonly used in Myanmar language.

```

Input: English noun phrase
Output: Tword //Translated Myanmar words
Begin
For each wordn
  If wordn is equal articles or cardinal numbers and n
  is not equal zero
    Tword += Translate wordn;
    Tword += CIA(wordn-1);
  Else
    Tword += Translate wordn;
End for
End
    
```

Fig.6. Translation Algorithm

Where

English noun phrase = (word₀, word₁,..., word_n) with Myanmar word order

n = no. of word in noun phrase

CIA = Classifier Identification Algorithm

```

Input: W // wordn-1
Output: Cs //Classifier
Begin
1. Extract Myanmar word and group of W from Bilingual Lexicon
2. Check last syllable of Myanmar word in particle list
  If Ls contains in Part
    Cs := Ls; break;
  Else
    go to step 3;
3. Check Group type.
  If Group for noun is "people" {
    If W equals "monk" or "nun" or "novice"
      Cs := "ပါး"; break;
    else
      Cs := "ယောင်"; break;
  }
  Else If Group for noun is "animal"
    Cs := "တောင်"; break;
  Else If Group for noun is "vehicle"
    Cs := "စီး"; break;
  Else
    go to step 4;
4. Check Subgroup
  For each Subgroup1
    If Myanmar word contains in Subgroup1
      N:= I;
  End for
  Switch(N)
    case 1: Cs := "ချောင်း"; break;
    case 2: Cs := "လုံး"; break;
    case 3: Cs := "ထည်"; break;
    case 4: Cs := "ချိတ်"; break;
    ...
    default: Cs := "ခု"; break;
return Cs;
End
    
```

Fig.7. Classifier Identification Algorithm (CIA)

Where

W = input English noun word

Part = set of particles list;

Ls = last syllable of word;

Group = people, animal, vehicle, things, other;

Subgroup = subgroup₁, subgroup₂, ... subgroup_i;

I = no. of subgroup;

N = number variable;

E. Target Sentence Generation

When the source sentence is translated to target sentence, different structures of the two languages is one of the difficulties. This problem is solved by using SCFG rules to get appropriate target sentence structure. After transforming structure, target words are specified by using bilingual lexicon. After finishing morphological synthesis, the final stage of machine translation system is target language generation. Target sentence corresponding with source sentence is generated by the system.

V. EVALUATION RESULT

In order to measure the performance of the system, Myanmar-English machine translation system is tested by using the example sentences. Testing set of example sentences consist of over 1200 sentences. In addition, this system used Myanmar-English bilingual lexicon which contains 13373 words to translate as Myanmar words. The evaluation measures for Myanmar-English machine translation system are defined in terms of the success rate, which is calculated in the following equation.

$$SuccessRate = \frac{No. \text{ of Translated Correct Sentences}}{Total \text{ No. of Sentences}} * 100\% \tag{1}$$

In English to Myanmar Machine translation, the sentence is determined as the correct sentence when not only all words in a sentence were translated but also the translated sentence was meaningful sentence. Although all words in a sentence were translated, if the translated sentence was not acceptable, it is incorrect sentence. For example, “I am glad to hear your good news.” is translated as “ငါသည် သင်၏ကောင်းသော သတင်းကြားရန်ဝမ်းသာသည်”, “This yard is full of children.” is translated as “ဤကိုက်သည် ကလေးများ၏ပြည့်စုံသည်”. In these sentences, although all words are translated, meaning of translated sentence is not acceptable. Some words are ambiguous words for verb such as “take”, “get”, “like” and “make” and for noun such as “leaves”, “yard” and “right”. In addition, translated sentence may be incorrect because of mistake of the parser. The word may be either verb or preposition and either adverb or preposition. Therefore, the Stanford parser may be wrong parsed tag. For example, the sentence “My father said me that I look like a queen.” is translated as “ငါ့အဖေသည် ငါ့ဘုရင်မတစ်ပါးကဲ့သို့ကြည့်သည်ဟု ငါ့ကိုပြောခဲ့သည်”. Moreover, when some words which is in

a source sentence did not include in lexicon, translation of such sentence is incorrect sentence. If translated sentence is lack of particles, it is determined as incorrect sentence. For example, the sentence “The stone was too heavy for him to lift it.” is translated as “ကျောက်တုံးသည် သူ့ငင်းမ ရန်အတွက်အလွန်လေးလံသည်” and so on. Table 3 describes evaluation of success rate for English to Myanmar machine translation on testing set containing around 1250 sentences.

In Myanmar to English machine Translation, the sentence is determined as the correct sentence when not only all words in a sentence were translated but also the translated sentence was meaningful sentence. Although all words in a sentence were translated, if the translated sentence was not acceptable, it is incorrect sentence. For example, the sentence “သူသည် သူမကို ကွန်ပျူတာ လက်ဆောင် ဝေးသည်” is translated as “He gives computer gift her.”, “မေမေသည်ထမင်းနှင့်ဟင်းကိုချက်သည်” is translated as “Mother cooks cooked rice and curry.” and “သူတို့သည် စာသင်ခန်းတွင်အလွန်လိမ္မာသည်” is translated as “They is clever very in classroom.”. In these sentences, although all words are translated, meaning of translated sentence is not acceptable. Sometime, translated sentence may be incorrect because of mistake of the parser. ML2KR’s parser cannot correctly parse some words in a sentence. For example, the sentence “သူတို့သည်ကမ်းခြေသို့ခရီးထွက်သည်” is translated as “they travels to leg shore.”. Some sentences are incorrect because of ambiguous words. For example, the sentence “ယနေ့ရာသီဥတုသည်အေးသည်” is translated as “Climate colds today.” and so on. Such these sentences are determined as incorrect sentences. The evaluation result of Myanmar to English system is shown in Table 4 on testing containing 1207 sentences. These sentences were tested on testbed specification that processor is Intel(R) Core(TM) i5-2430M CPU @ 2.40GHz 2.40 GHz and RAM is 4.00GB and Operating System is 32-bit OS.

Table 3. Evaluation Result of Myanmar to English Machine Translation

E to M System	No. of Sentence	Success Rate	Average Translation Time
Correct Sentence	1030	82.14%	3 sec
Incorrect Sentence	224		
Total	1254		

Table 4. Evaluation Result of English to Myanmar Machine Translation

M to E System	No. of Sentence	Success Rate	Average Translation Time
Correct Sentence	971	80.45%	8 sec
Incorrect Sentence	236		
Total	1207		

VI. CONCLUSION

Bidirectional translation system is implemented for Myanmar and English language using rule based machine translation. Therefore, SCFG rules are generated via CFG

rules according to the respective parser. There are 129 SCFG rules for English to Myanmar tree transformation by using Stanford Parser. Moreover, there are 800 SCFG rules for Myanmar to English tree transformation by using ML2KR's parser. Total number of words in Myanmar-English bilingual lexicon is 13373 words. This system can support for Myanmar-English language communication. It can get acceptable result and can be applied in real world. When over 1200 simple sentences are tested, success rate is 80% of Myanmar to English Translation and 82% of English to Myanmar Translation.

REFERENCES

- [1] P. J. Antony, "Machine Translation Approaches and Survey for Indian Languages" The Association for Computational Linguistics and Chinese Language Processing Vol. 18, No. 1, March 2013, pp. 47-78.
- [2] S.R.Priyanga, AP, A.AzhaguSindhu, AP, "Rule Based Statistical Hybrid Machine Translation", International Journal of Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-1, Issue-5, April 2013.
- [3] Khaled Shaalan et.al, "An English-Arabic Bi-directional Machine Translation Tool in the Agriculture Domain", IFIP International Federation for Information Processing 2010.
- [4] Mr.Uday C. Patkar et.al, "Transformation of Multiple English Text Sentences To Vocal Sanskrit Using Rule Based Technique", International Journal of Computers and Distributed Systems, Vol. No.2, Issue 1, December 2012.
- [5] Fai Wong et.al, "Handheld Machine Translation System Based on Constraint Synchronous Grammar", Machine Translation Summit XIII, Sep. 2011, p. 439-446.
- [6] Shibli Syeed Ashrafi et.al, "English to BanglaMachine Translation System Using Context-Free Grammars", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 2, May 2013.
- [7] T. T. Zin et al, "Myanmar Phrases Translation Model with Morphological Analysis for Statistical Myanmar to English Translation System", International Journal of Computer Applications, Volume 28, No 1, 2011, pp 13-19
- [8] http://en.wikipedia.org/wiki/Statistical_machine_translation
- [9] <http://nlp.stanford.edu/software/lex-parser.shtml>
- [10] S.L. Phye, "Development Of Lexico-Conceptual Knowledge Resources And Syntax Analyzer For Myanmar Language", Ph.D Thesis, University of Computer Studies, Mandalay, 2013.
- [11] R. Harsha Wardhan, "Rule Based Machine Translation System for English to Malayalam Language", Centre for Excellence in Computational Engineering and Networking, December, 2011
- [12] A. Lopez, "Statistical machine translation", ACM Computing Surveys (CSUR), 40(3), p.8, 2008.
- [13] T.W. Mamta, "Survey of Approaches Used in Machine Translation System", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 5, May 2015
- [14] R.S. Narayanan, "English-Telugu Rule Based Machine Translation system", 2012.
- [15] S. Pal, "Improved Alignment in Phrase Based Statistical Machine Translation System." PhD diss., Jadavpur University, Kolkata, 2013.

Authors' Profiles



Yin Yin Win is a Ph.D-candidate at the University of Computer Studies, Mandalay, Myanmar. She got B.C.Sc (Hons) degree and M.C.Sc degree from Computer University, Mandalay. Her research interests mainly focus on Natural Language Processing, especially in Machine Translation.



Aye Thida is an Associate Professor, at Faculty of Computer Science department (Artificial Intelligence Lab), University of Computer Studies, Mandalay (UCSM), Myanmar. She was a leader of Natural Language Processing Project. Her team has developed Myanmar to English Translation System in 2011. Her research interests include Distributed Processing, Queuing and Natural Language Processing. She is currently working NLP researches. Dr. Aye Thida received B.Sc(Hons) Maths degree from the Mandalay University, Myanmar and her M.I.Sc and Ph.D degrees in Computer Science from the University of Computer Studies, Yangon(UCSY), Myanmar.

How to cite this paper: Yin Yin Win, Aye Thida, "Myanmar-English Bidirectional Machine Translation System with Numerical Particles Identification", International Journal of Information Technology and Computer Science(IJTCS), Vol.8, No.6, pp.37-43, 2016. DOI: 10.5815/ijitcs.2016.06.05