ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 7, September 2012

Morphological Processor for Inflectional Case of Multipurpose Lexico-Conceptual Knowledge Resource

Soe Lai Phyue, Aye Thida

***Abstract*— Myanmar language is morphologically rich and agglutinative language. Myanmar words are postpositionally inflected with various grammatical features which can cause difficulties for Language Acquisition (LA). LA is important for development of Myanmar Natural Language Processing (NLP). For this reason, Myanmar Language Lexico-Conceptual Knowledge Resource (ML2KR) is developed. ML2KR consists of three main resources; Myanmar WordNet, Myaanmar-English Computational lexicon and Morphocon. Myanmar WordNet and Myanmar-English computational Lexicon serve as Myanmar Language Resources (MLRs), and Morphocon serves as morphological processor to support the MLRs. This paper presents the morphological processor (analyzer and generator), Morphocon, to support the inflectional verbal and colloquial cases for knowledge resources by using the rule-and-feature based model of Myanmar inflectional morphology. By supporting with Morphocon in MLRs, it can reduce the time and storage consumption. The evaluation of coverage for lexical acquisition increased to nearly tenfold of existing data. Moreover, the evaluation of the correctness of Morphocon yields the satisfactory result because precision, recall and f-measure are nearly and over 95% in both morphological analyzer and generator. Error analyzes of morphological analyzer shown that it need to deal for complicated word of inflectional case for negative of verb and superlative of adjective.**

***Index Terms*— Lexical Acquation, Natural Language Processing, Myanmar Language Lexico-Conceptual Knowledge Resource, Myanmar Language Resources, WordNet.**

I. INTRODUCTION

Natural Language Processing (NLP) is the one of the essential research for many productivity tools in many computer applications. The development of language resources for Myanmar and its availability is a must for enhancing language processing capabilities. Morphological processor which is of increasingly great significance for Machine Translation (MT), Information Retrieval (IR) and Lexical Acquisition (LA), is needed to handle the inflectional case not only for a bilingual concept lexicon but also for monolingual lexicon [7].

Manuscript received Sep 15, 2012.

***Soe Lai Phyue**, Ph.D Candidate, University of Computer Studies, Mandalay, Myanmar.*

Aye Thida, Research and Development Department, University of Computer Studies, Mandalay,Mandalay, Myanmar.

Myanmar Language is official language for The Republic of the Union of Myanmar and there is no doubt in the necessity of constructing basic language processing resources for it. Beside then, it also needs to construct Myanmar language analyzer and target language generator to overcome the limitation of the developing in Myanmar NLP applications. For these requirements, we have been developed the Myanmar WordNet and computational bilingual computational lexicon as MLRs [9].

In the previous work, MLR is constructed by using semiautomatic methodology by acquiring the lexical and conceptual knowledge from WordNet and Myanmar<->English Machine Readable Dictionaries (MRDs). To build the MLRs, the translation links are collected from existing bilingual MRDs and semantic meaning and synset links are collected from English WordNet. The collected links and their meaning are manually verified. The computational lexicon stores the word according to their part of speech. However, this work needs to deal the inflectional cases for MLRs to improve the coverage. Through a detailed study of the Myanmar language, we have been able to develop an analyzer that incorporates many of the unique features and challenges present in Myanmar. Although the analyzer of Myanmar Language is sufficient for Myanmar WordNet, Myanmar English lexicon is needed to translate according to their inflected case by generating related word.

This paper considers morphological processor which is called Morphocon for MLRs. Morphological processor is intended to provide an analysis as well as a generation for every derived item of the Myanmar language resources. Although morphological processor covers both inflectional and derivational morphology, Morphocon of ML2KR is currently treated only for inflectional case of noun, verb and adjective. Therefore, Morphocon for MLRs consists of (i) determining its stem by a morphological analysis and (ii) generating all or a subset of the permissible word forms.

This paper is organized as follows. The similar works of morphological processor for other languages are introduced in Section 2. The background information of language resources methodology is described in Section 3. Section 4 sketches the morphological processing between Myanmar and English word. Statistical result of Morphocon and coverage of MLRs due to Morphocon is expressed in Section 5. Finally, Section 6 draws the conclusion and future work.

II. Related Work

Due to the difficulties of language acquisition in morphological rich language, the development in NLP applications have limitation. To overcome this problem, morphological analyzer for source language and generator for target language, called morphological processor, are considered. It became the essential part of every NLP applications. In 2006, Thai Phuong Nguyen and Akira Shimazu [11] proposed morphological transformational rules and Bayes' formula based transformational model to translate English to Vietnamese.

A morphological processor for Modern Greek is presented in [1]. The morphological processing is controlled by a finite automaton and it combines a dictionary containing the stems for a representative fragment of Modern Greek and all the inflectional affixes with and a grammar which carries out the transmission of the linguistic information needed for the processing. The words are structured by concatenating a stem with an inflectional part. In certain cases, phonological rules are added to the grammar in order to capture lexical phonological phenomena.

A finite transducer that processes Spanish inflectional and derivational morphology is presented in [3]. The system handles both generation and analysis of tens of millions inflected forms. Lexical and surface (orthographic) representations of the words are linked by a program that interprets a finite directed graph whose arcs are labeled by n-tuples of strings.

Morphology generation models of English to Russian and Arabic for machine translation are presented in [6]. They applied their inflection generation models in translating English into two morphologically complex languages, Russian and Arabic and their model improves the quality of SMT over both phrasal and syntax-based SMT systems according to BLEU and human judgments.

In [12], they presented the two-level framework, as it is well known; morphographemics and morphotactics. Morphographemics is modeled in two level rules (TLR) and morphotactics either in continuation classes or in unification word grammars (WG). In their system, 114 rules cover nominal infection and 10 rules cover verbal infection.

Myanmar language is rich morphology; it has very limitation in development of NLP application. The best way to cover the morphological changing is using finite state automaton (FSA). This method is sample, easy to implement and improve the correctness for Morphocon. Therefore we implement the morphological analyzer for Myanmar word and generator for relevant English word using rule and feature based FSA. In this work, the evaluation result on sentence is over 95% of precision, recall and f-measure although statistical method not being used.

III. MM2KR Framework

The architecture of ML2KR is multipurpose in the sense that it is multifunctional. Thus, it has been designed to be potentially reused in many NLP tasks such as Lexical Analyzer, Myanmar to English Translation system, Machine Readable Dictionary. ML2KR consists of three major resources which are Morphocon, Myanmar WordNet and bilingual computational lexicon as shown in Figure. 1. These resources are formed several independent but interrelated modules.

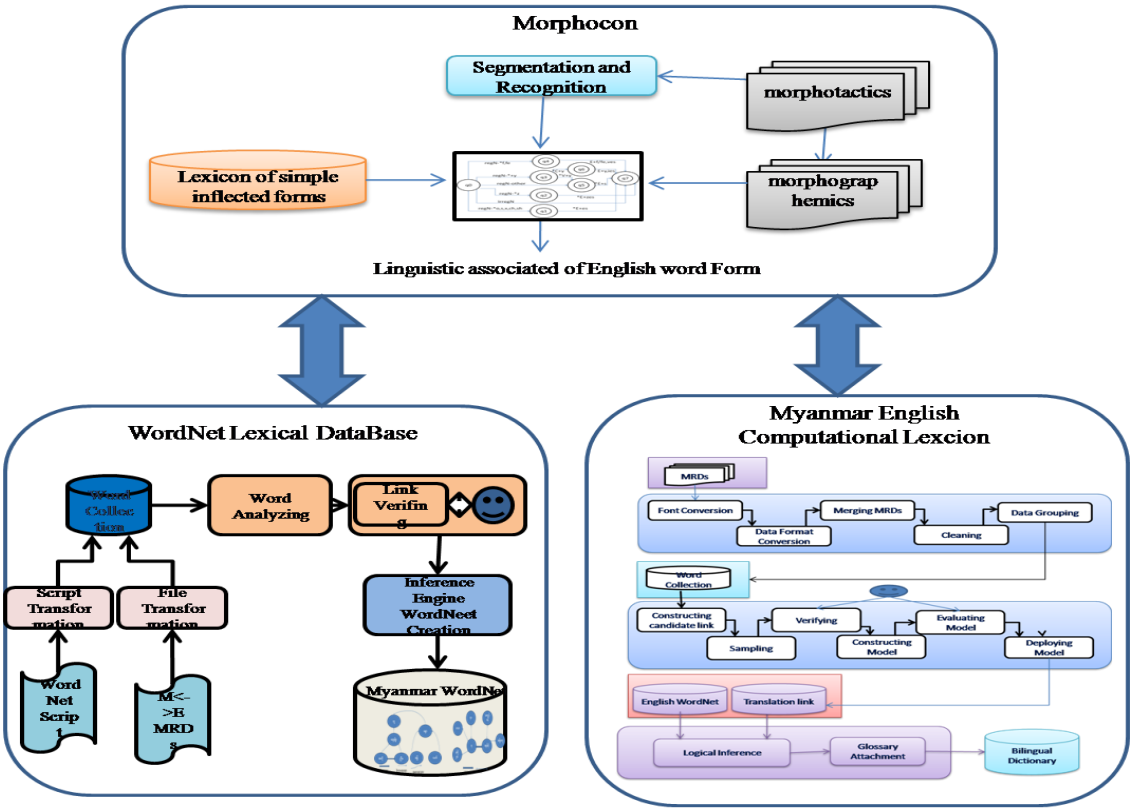

Figure.1. ML2KR Framework

A. Myanmar WordNet

The most ambitious feature of Myanmar WordNet, however, is its attempt to organize lexical information in terms of word meanings, rather than word forms. Myanmar WordNet which contained semantic concept is crucial role in NLP because WordNet gives not only lexical information of words but also semantic meaning of each word. Manual construction of WordNet is the most reliable technique for obtaining structured lexicons but it is costly and highly time-consuming [4]. Myanmar WordNet is constructed by using the origin of WordNet lexical database and Myanmar<->English MRDs with the semiautomatic methodology. The relations WordNet applied to the noun and verb concepts are synonymy, antonymy, hypernymy, holonymy, entailment, cause and etc., among which synonymy and hypernymy are the most important. Synonymy and hypernymy help to form the SynSets and their hierarchies respectively. The hypernymy tree, as the hierarchy of Concepts, provides a common way of making induction for the NLP researchers [2]. According to the specification of WordNet, the noun is categorized into 26 broad, the verb is categorized into 15 broad and the adjective is 3 broad. The words in Myanmar WordNet are stored as stem words and it does not contain in inflected form of noun as plural and verb of various tense forms. Myanmar WordNet covered with 173853 links, 38142 synsets and 25865 words.

B. Myanmar English Bilingual Computational Lexicon

Computational lexicons are among the most important resources for natural language processing (NLP). Their importance is even greater in languages with rich morphology, where the lexicon is expected to provide morphological analyzers with enough information to enable them to correctly process intricately inflected forms [8]. Every lexicon item belongs to a part of speech category. The part of speech of an entry determines its additional attributes. Bilingual lexicon structure is mostly depending on Myanmar WordNet. The Bilingual computational lexicon is built together with Myanmar WordNet construction. Verbs are specified for their function and their class according to the WordNet structure. Bilingual computational lexicon has total 71 tag set which consists tag sets: 26 for noun, 15 for verb, 3 for adjective, 2 for adverb, 17 for proposition and 8 for conjunction. But it also lists several lexical properties which are specifically targeted at morphological analysis. The bilingual lexicon stores two main pieces of information: a root and an inflection pattern.

C. Morphocon

As An important class of lexical relations are the morphological relations between word forms. Computational lexicon and WordNet as a language resource became increasingly obvious that have to deal with inflectional morphology [12]. For example in WordNet, the word of input sentence or phrase include as “ ” and clicked a request for information, WordNet should not reply that the word was not in the database. A program was needed to strip off the plural suffix and then to look up “ ” which certainly is in the database. On the side of lexicon, we need to generate the relevant information as “trees” in English word. This need led to the development of a program for dealing with inflectional morphology. Although the inflectional morphology of Myanmar is relatively simple, writing a computer program to deal with it proved to be a more complex task than had been expected.

This paper considers morphological analysis on number category of noun, suffixes and prefix particle of Myanmar verb and degree of adjective. Plural number particles of Myanmar noun are “ ”. Myanmar verb has many particles and suffixes. It is not easy to define tense like English. Some suffixes have same meaning. (i) For example: these suffixes: “ ” have same meaning. Some verb behave particle to support previous verb in sentence. (ii) For example: “ပြောပေးသည်: talk” in this verb “ ” behave particle to support previous verb “ပြော”. But “ပြော: *talk*” and “ : *give*” can behave individual verb in sentence. More than two individual verbs can include in Myanmar compound verb. (iii) For example: “ဝင်ထွက်သွားလာသည်”- four individual verb “ဝင်+ထွက်+သွား+လာ” includes in this compound verb. It is difficult to translate English language. Some verb particles indicate tense of Myanmar verb. (iv) For example: (သည်: present tense, : past tense, : continuous tense, မည်၊ လိမ့်: future tense). This paper focuses on singular and plural number of noun, verb suffixes which have same meaning in translation and particles which indicate verb tense and adjective degree [14]. Therefore, an important class of lexical relations is the morphological relations between word forms.

IV. Development Strategies of Myanmar Morphology

Myanmar Language is known for its morphological richness and complexity. Myanmar morphology has always been a challenge for computational processing and a hard testing ground for morphological analysis technologies. There are mainly two strategies for the development of Myanmar morphologies depending on the level of analysis:

1. Stem-based morphologies: analyzing Myanmar at the stem level and using regular concatenation. A stem is the least marked form of a word that is the uninflected word without suffixes, prefixes, proclitics or enclitics. In Myanmar, this is usually the perfective, 3rd person, forms of verb, and in the case of nouns and adjectives they are in the indefinite form.
2. Root-based morphologies: analyzing Myanmar words as composed of roots and patterns in addition to concatenations. A root is a sequence of three (rarely two or four) consonants which are called *radicals*, and the pattern is a template of vowels, or a combination of consonants and vowels, with slots into which the radicals of the root are inserted. This process of insertion is usually called *interdigitation*. With this technique, morphological structure of words also can be analyzed.

In our implementation we adopted the idea that a root is an abstract form that does not belong to a specific POS, but it plays a crucial part in stem formation. So using the stem as base form is far less complex in developing and maintaining, less ambiguous, and more suitable for syntactic parsers that aim at translation.

V. Propose Morphological Processor (Morphocon)

The quality of a morphological analyzer is greatly depends on the quality of the lexicon. A morphological analyzer must consult with the lexicon to check whether a theoretical analysis of a word indeed belongs to the language. This system facilitates a modular development of morphological analysis and disambiguation systems. The morphological analyzer interacts with, but is separated from the lexicon. Proposed Morphocon performs analyzing Myanmar words and generating the equivalent English words: this is basically the rule of the morpheme of the Myanmar word for WordNet and grammar pattern relation between Myanmar and English word for lexicon.

The basic idea is to generate all the inflected forms Myanmar word to English word which is useful for NLP application and induced by the lexicon. It is common to think that for languages with rich morphology such a method is impractical. While this may have been the case in the past, contemporary computers can efficiently store and retrieve millions of inflected forms. Of course, proposed Morphocon would break in the face of an infinite lexicon (which can easily be represented with FST), but for most practical purposes it is safe to assume that natural language lexicons are finite. The morphological analyzer is obtained by inflecting the base forms in the lexicon. The numbers of inflected forms are used by the analysis program and generate the equivalent Myanmar word. The framework of morphological processor is shown in Figure. 2.

The analysis start with application of a decomposition system defined by morphological grammar, to each word of the text to identify it's radical and affixes. In the second step, grammars (finite-state transducers) produce lexical constraints checking the validity of segmentation thanks to a dictionary lookup. So, these grammars associate the recognition of a word to lexical constraints, working only with valid combinations of the various components of the form. Typically there are several output strings, each representing a possible analysis of the input word.

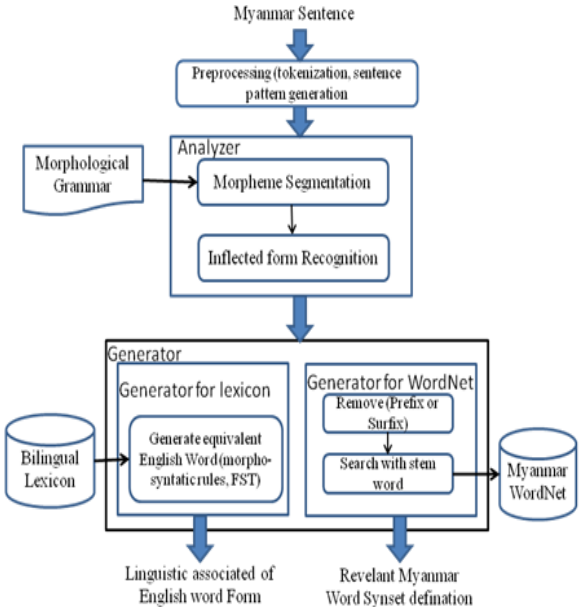

Figure 2. A Framework for Morphological Processor

A. Preprocessing of Input Sentence

Preprocessing step includes tokenization, segmentation and pattern merging for input sentence which are present in [10]. To process text computationally, syllables have to be determined first. Since Myanmar linguistic tradition there is not a clear-cut, we need to pass the tokenized syllable as one of the steps. Word tokenization is done by rules based approach.

In linguistics, a word is a basic unit of language that carries meaning and can be spoken or written. It can consist of one or more morphemes that are linked more or less tightly together. Typically, a word will consist of a root or stem and zero or more affixes. Without a word segmentation solution, no NLP application (such as Part-of-Speech (POS) tagging and translation) can be developed. Words can be combined to form phrases, clauses and sentences. Myanmar word may consist of two or more stems and joined together is known as a compound word.

B. Morphological Analyzer

Morphological Analysis as a vocabulary acquisition strategy has both its advocates and antagonists. Since most Myanmar words consist of a stem, which mainly specifies the lemma, and a set of affixes that mainly specify the morphosyntactic features, it is appropriate to concatenate the prefix or suffix elements. We take a somewhat profit of nature of Myanmar language to defining and computing word relations to its application in a morphological processor for Myanmar morphological analyzer. The main advantage of this process is the extreme simplicity both of its tagging process and of their interpretation.

Therefore, the analyzer of Morphocon consists of two processes. The first is to define the stem by segmenting with prefix and suffix and the second is recognized the inflected form of stem by their prefix and suffix.

C. Morphological Generator

The generator of Morphocon is apart from two cases; Myanmar WordNet and bilingual lexicon. Myanmar WordNet, useful in word sense disambiguation, does not need to translate the equivalent English word and it only need to remove irrelevant word as synset. For example, if we look up in "□□□□□□□□□ " in Myanmar WordNet, we retrieve the information of "□□□□ " as noun. For the bilingual lexicon, we need to define their case and translate as the equivalent English meaning as "teeth". The implementation is based on the concept of validation grammars. The morphological processing is controlled by a finite automaton. The detail study of generating process for each POS is described in following section.

VI. Finite State Automaton for Morphological Generator

A. Relationship between Myanmar and English Noun Form

The noun in Myanmar word can have a suffix indicating plurality. It can be pluralized by suffixing the particle "□□□ " in colloquial Myanmar or "□□□□ " in formal Myanmar. The

particle "□□□ " which indicates a group of persons or things, is also suffixed to the modified noun. . To generate plural or singular forms of English word, we use English grammar rules. Singular words which end is s, z, sh, ch or x, we add es to become plural words. Singular words which end is consonant with "y" changes the "y" to "i" and add es. All other singular words add "s". But some nouns have irregular form e.g; man (plural men). We cannot handle this irregular noun. There we used the Finite State Automata (FSA) as followed in Figure 3.

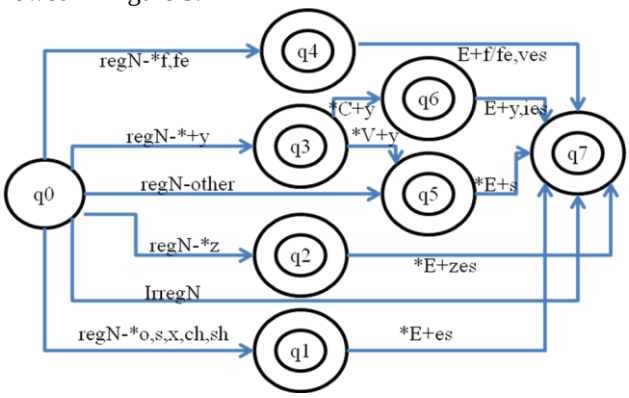

Figure 3. Finite state automata for Singular noun to Plural Form

B. Relationship between Myanmar and English Verb Form

In the affirmative, the order of elements is V [one or more roots, possibly compounded] (+auxiliary verb) + aspect particle + modal ending. The most commonly used verb particles and their usage are shown below with an example verb root "□□□ ". Alone, the statement "□□□ " is imperative. The suffix "□□□ " (literary form: □□□) can be viewed as a particle marking the present tense and/or a factual statement. The suffix □□□ denotes that the action took place in the past. Note that the suffix "□□□ " in this case denotes a factual statement rather than the present tense. We also generate verb tense by using verb stem word and suffixes particles. Stem of verb add "ed" to become past tense. We use English grammar rule to change verb tense but some verb has irregular form e.g; past tense of "read" is also "read". We handle irregular verb by using irregular verb list defined by Oxford Dictionary. The particle "□□ " is used to denote an action in progression. It is equivalent to the English '-ing'". This particle "□□□ " which is used when an action that had been expected to be performed by the subject is now finally being performed, has no equivalent in English. So in the above example, if someone had been expecting you to eat and you have finally started eating, the particle □□□ is used. The particle "□□□ "□□□ "□□□□□□□ "□□□□□□□ " are used to indicate the future tense or an action which is yet to be performed. An FSA for English derivational morphology structure for English word is as shown in Figure 4.

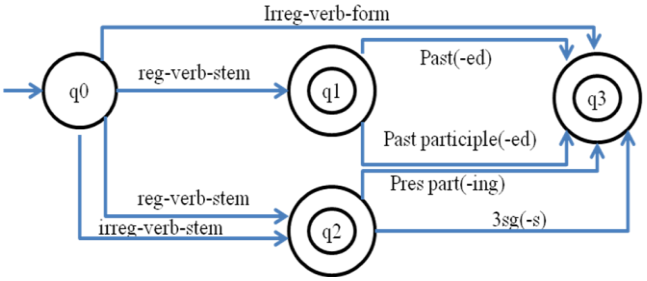

Figure 4. Finite state automata for Verb Tenses

C. Relationship between Myanmar and English Adjective Form

In Myanmar word, adjective is defined by the word with suffix "□□□ ","□□□□ ","□□□□ " Beside then, it has verbs that carry the meaning "to be X", where X is an English adjective. Comparatives are usually ordered: X + "□□□□□□□ " "□□□ " "□□□□ " + adjective, where X is the object being compared to. For this case , we add the X word in adding more as prefix or suffix as X-er. Superlatives are indicated with the prefix □ + adjective + □□□□ . In figure 5, degree of adjective transformation is shown.

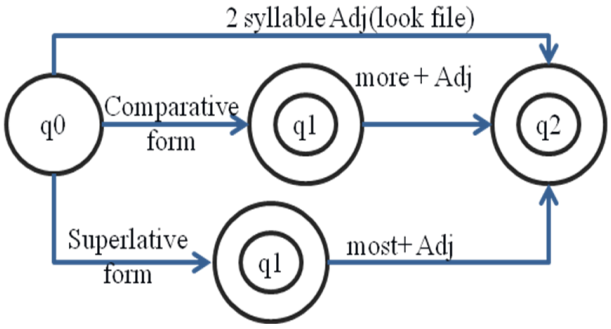

Figure 5.Finite state automata for degree of adjective

VII. Evaluation and Statistical Result

A. Statistical improvement of the Computational Lexicon

If the lexicon and Myanmar wordNet used as machine readable dictionaries, Morphocon can help for rich vocabulary. To acquire the inflectional case of Myanmar word, we must it control with morphological processor.

The lexicon use the following parts of speech: noun (N), verb (V), pronoun (Pron), adverb (Adv), Adjective (Adj), preposition (Prep) and conjunction (Conj). Table 1 shows the distribution of these parts of speech in the two formats: the first column is the distribution of the root forms in the lexicon files, and the second column is tile distribution for the inflected forms derived from the Morphocon. As a result we have covered with 3 fold increased for noun except for group of noun, 12 fold for verb and 3 fold for adjective in lexicon for Myanmar word. According to the result, proposed Morphocon yielded a satisfactory result for the reduction of the storage of lexicon and time consuming for the entry of inflected word between Myanmar and English words.

Table 1.Statistic of the Computational Lexicon

parts of speech	No. of root form	No. of inflected form
Noun	31243	110846
Verb	12720	152640
Pronoun	170	170
Adverb	3369	3369
Adjective	6819	20457
Preposition	108	108
Conjunction	185	185

B. *Evaluation of the System*

The constructed morphological processor is evaluated using the well-known measures precision, recall, and the F-measure in equation 1, 2 and 3. In this study, we adopt the Myanmar Word Segmentation and translation based on the Myanmar WordNet and Myanmar-English WordNet like lexicon is experimented [10].

In [10] morphological analyzer is used to cover the acquisition of vocabulary and tag for subjective verb agreement in Myanmar to English translation system. The processor is used to translatie the equavilent words of Myanmar to English words. Myanmar word can define as verb in one word (eg. "□□□□ □□□ □□□ ") and some word defined as verb and particle in two word (eg. "□□□□ " and "□□□ □□□ "). In that case, Morphocological generator has to gerneate the one word (eg. Went) whatever it take as one or two word in Myanmar word. The Morphological processor of inflectional case is used in the Part of Speech Tagging (POST) process.

We test the system in general domain. Sentence types in testing case are simple and compound. The length of source sentences consists of word between 5 and 15. Only single references are used in this measure. These reference sentences are manually translated. This system does not consider word order of Myanmar and English language. Therefore, we ignore the word order of candidate and reference sentences.

$$\text{Precision } (C \mid R) = \frac{|C \cap R|}{C} \quad (1)$$

$$\text{Recall } (C \mid R) = \frac{|C \cap R|}{R} \quad (2)$$

$$F - measure = \frac{2*(precision*recall)}{(precision+recall)} \quad (3)$$

C=set of candidate sentences
R=set of reference sentences

As this morphological processor only considers for adjective, noun and verb morphology, we have limitation in analysis process. We tested with 100 sentences in which word lengths are between 5 and 15. We use Myanmar3 font. In proposed morphological analyzer, the precision is 97.8%, recall is 98.34% and F-measure is 98.06%. In the generation process, the precision is 95.67%, recall is 92.42% and F-measure is 94.01% followed by analyzer. The precision, recall and F-measure of generation are less than analyzer because the error of analyzer effect on generation process which are shown in Figure 6.

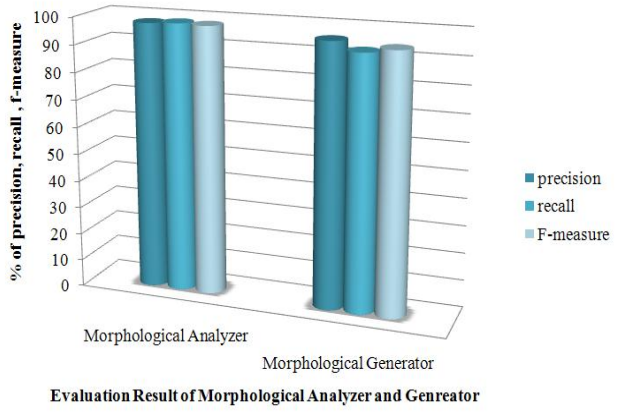

Figure 6. Evaluation Result of Morphological Processor

C. *Error Analysis of Morphocon*

Noun has error in generator of Morphocon, so error of noun does not effect on Myanmar WordNet. It has no error occured in analyzer. The generation of some equivalent Myanmar to English word in plural form has informal case. It also depends on the wordlist of lexical term of informal plural case. Errors in proposed system are as follow. Compound verb (□□□□□□□□□□□□) has two meaning. (□□□□□□) and (□□□□□□□□□) and meaning of (□□□□□□□□□□□□□□) is (went and ate). Although proposed Morphocon of generator can translate it as (□□□□□□□ : go) and (□□□□□□□□□ : ate), we have difficulty to translate (□□□□□□□□□□□□□ : went and ate) to get correct translation. Some verbs support to previous verb ("□□□□□□□□□□ ": give), correct translation is "talk". Beside then in the negative inflection of verb has more error because negative particle of Myanmar "□" can take as prefix or middle of stem verb such as ("□□□□□□□□□ ":not tell) and ("□□□□□□□□□□□□□ ":not listen). In the latter case ("□□□□□□□□□□□□□ "is analyzes as "□□□ " and "□□□□□□□□□□ " which as (ear and not stand).

In adjective, we have same error like negative verb inflection like ("□□□□□□ ": respectful) of negative form as ("□□□□□□□ ": not respectful) or ("□□□□□□□□ ": not respectful). Although the word of "□□□□□□□ " is not a problem in analyzer, the word "□□□□□□□□ " has error. Another error in adjective ("□□□□□□□□□□□□□ ": good looking) is superlative degree for compound adjective ("□□□□□□□□□□□□□□□□□□□ ": best looking). In this word, the analyzer split words as "□□□□□ " and "□□□□□□□□□□□□□□□ ".

Figure 7 shows the error analysis for Morphocon in their POS. This chart depends on the evaluation of Morphocon error. According the analysis, the verb has more error than other POS because it has very complicated word form for inflection. In adjective, main error

occurs in superlative degree and negation. Noun is simple and less inflected word than other POS.

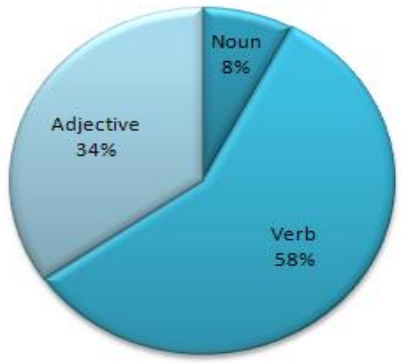

Figure 7.Error Analysis for Morphological Processor According their POS

VIII. CONCLUSION

We presented the ML2KR frame work which consists of Morphological processor, Myanmar WordNet and bilingual Lexicon. Proposed Morphological processor is independent of ML2KR. In this work, we proposed Morphological processor and we emphasize on inflectional case of Myanmar morphology. The contributions of proposed rule based Morphocon include two fold: morphological analyzer for Myanmar words and morphological generator for translation of Myanmar word to English word by using FSA. According to the evaluation results, proposed morphological processor can improve the lexical acquisition of WordNet and bilingual lexicon (MLRs) and it can be used in further NLP applications. We also found that rule based morphological analyzer has errors in inflectional Myanmar word form and it can be overcome by applying corpus based statistical approach. As a future work, we will apply corpus based statistical approach in Morphological Analyzer and To become a complete Morphological processor, we will consider derivational case of Myanmar morphology.

REFERENCES

[1] A. Ralli, E.Galiotou, A Morphological Processor for Modern Greek, Proceedings of the third conference on European chapter of the Association for Computational Linguistics,1987, Pp.26-31

[2] C. Fellbaum, WordNet: An electronic lexical database. MIT Press, Cambridge, Massachusetts, 1998.

[3] E.Tzoukermann and M. Y. Liberman, A Finite State Morphological Processor for Spanish, Proceedings of the 13th International Conference on Computational Linguistics (COLING 90), 1990, Pp. 277-281

[4] G. Prószéky and M. Mihăltz. Semi-automatic Development of the Hungarian WordNet, Proceedings of the LREC 2002 Workshop on WordNet Structures and Standardization, Las Palmas, 2002.

[5] K.Daniel, S.Yve, Z. Martin, E.Dania, A Freely Available Wide Coverage Morphological Analyzer for English, Proceedings of COLING-92, NANTES, 1992, Pp. 950-954

[6] K. Toutanova, H. Suzuki, and Ruopp, Applying morphology generation models to machine translation, Proceedings of Association for Computational Linguistics: HLT. Columbus, Ohio, 2008, Pp.514-522.

[7] M. Neff et. al, Get It Where You Can: Acquiring and Maintaining Bilingual Lexicons for Machine Translation, Proceeding of AAAI, 1993.

[8] O. Dzikovska, D. Swift, F. Allen, Building a computational lexicon and ontology with FrameNet, Proceeding of the Workshop on Building Lexical Resources from Semantically Annotated Corpora, LREC 2004.

[9] S.L.Phyue, Construction of Myanmar WordNet Lexical Database, IEEE student Conference on Research and Development, Malaysia, 2011.

[10] S.L.Phyue, Lexical Analyzer for Myanmar Language, Proceedings of the International Conference on Computer Applications, Yangon, 2012, Pp.95-101

[11] T. P. Nguyen and A. Shimazu, Improving Phrase-Based SMT with Morpho-Syntactic Analysis and Transformation, Proceedings of the 7th Conference of the Association for Machine Translation in the Americas, Cambridge, August 2006, Pp. 138-147.

[12] T.Toni, The derivation of a large computational lexicon for a two level morphological analyzer for Catalan from a MRD, 1997.

[13] www.FunGramKB.com/architecture.htm

[14] Department of the Myanmar Language Commission,Ministry of Education, Union of Myanmar, "Myanmar Grammar", 2005