Myanmar Number Normalization for Text-to-Speech

Aye Mya Hlaing, Win Pa Pa
Natural Language Processing Lab.,
University of Computer Studies, Yangon,
Yangon, Myanmar
{ayemyahlaing, winpapa}@ucsy.edu.mm

Ye Kyaw Thu
Artificial Intelligence Lab.,
Okayama Prefectural University,
Okayama, Japan
ye@c.oka-pu.ac.jp

***Abstract*---Text Normalization is an essential module for Text-to-Speech (TTS) system as TTS systems need to work on real text. This paper describes Myanmar number normalization designed for Myanmar Text-to-Speech system. Semiotic classes for Myanmar language are identified by the study of Myanmar text corpus and Weighted Finite State Transducers (WFST) based Myanmar number normalization is implemented. Number suffixes and prefixes are also applied for token classification and finally, post-processing has been done for tokens that cannot be classified. This approach achieves average tag accuracy of 93.5% for classification phase and average Word Error Rate (WER) 0.95% for overall performance which is 5.65% lower than rule-based system. The results show that this approach can be used in Myanmar TTS system and to our knowledge, this is the first published work of Myanmar number normalization system designed for Myanmar TTS system.**

***Keywords*-Myanmar Number Normalization; Text Normalization; Weighted Finite State Transducer; Myanmar Text-to-Speech; Myanmar;**

I. Introduction

Text-to-Speech (TTS) synthesis is a technique for generating intelligible, natural-sounding artificial speech for a given input text. Typical TTS systems have two main components, text analysis and waveform synthesis [1]. Text normalization, phonetic analysis and prosodic analysis are needed to be done in text analysis components. Text normalization is the first and crucial phase of text analysis.

TTS systems need to work on real text and it contains non-standard words (NSWs), including numbers, dates, currency amounts, abbreviations and acronyms. NSWs cannot be found in a dictionary or cannot get their pronunciations by "letter-to-sound" rules [2].

Myanmar text contains many NSWs with numbers. If normalization of these NSWs is not performed in the initial step of Myanmar TTS system, the quality of the system will be degraded. Therefore, normalization of NSWs with Myanmar numbers are more emphasized for this research.

Most text normalization still involves hand-constructed grammars because statistical techniques typically require at least some annotated data to train models, which is often not available for all of the ambiguities one would like to resolve [3].

Some information of rule-based text normalization for Myanmar language are described in [4]. Rule-based and lookup dictionary based methods are applied in text normalization of Croatian [5], Bengali [6] languages. Finite State Automata (FSA) and Maximum Entropy (ME) Classifier and Rules are used as the text normalization strategy of Mandarin [7] language. Application of decision tree and decision list are explored in Hindi text normalization [8]. Language model, supervised and unsupervised approaches are applied in Russian number names [9] and Vietnamese text normalization [10]. Weighted finite-state transducers (WFSTs) are applied for the Kestrel text normalization system, a component of the Google text-to-speech synthesis system [3]. Recurrent neural networks (RNN) are applied to learn the correct normalization function from a large corpus of written text aligned to its normalized spoken form. Although RNN produces good results on overall accuracy, it is prone to make the occasional errors [11].

For Myanmar language, statistical approaches cannot be used on number normalization because there is no annotated or parallel normalized corpus for Myanmar number normalization. Therefore, Myanmar number normalization is implemented by writing number normalization grammars that are compiled into libraries of WFST.

This paper describes the semiotic classes for Myanmar language and the development of number normalization for Myanmar text-to-speech system by implementing two phases: classification and verbalization phases. Word segmentation process is needed as the pre-processing step because Myanmar text lacks white space between words. Language specific grammars based on WFST are applied for this system.

II. Defining Semiotic Classes for Myanmar Language

The first part in number normalization is classification of NSWs and the second part is verbalizing the detected NSWs into their standard words. For NSWs classification, the semiotic classes for Myanmar language is needed to identify. Based on the investigation of Myanmar sentences from Asian Language Treebank (ALT) parallel corpus[1] [12] and some Web data (3,150 sentences), we identified a set of semiotic classes for Myanmar Language. ALT corpus is the parallel corpus for seven languages in the news domain and comprises 20,000 sentences. The result semiotic classes are shown in Table I. NSWs (Myanmar digits) are shown in bold style in tables and contents of the paper.

Table I
Semiotic Classes for Myanmar

Semiotic Class	Description	Example:
DATE	date	**၁၂-၂-၂၀၁၇**, **၂၀၁၇** ဖေဖော်ဝါရီ **၁၂**, etc.
TIME	time	**၁၀း၂၀း၂၅**, **၁၀** နာရီ **၂၀** မိနစ်, ဂျီအမ်တီ **၀၈၃၀**, etc.
CURRENCY	currency amount	**၁၂၃,၄၅၀** ကျပ်, $**၃,၀၀၀**, etc.
NUMBER	cardinal, decimal	**၁၀** ယောက်, **၁၂၅.၄၅**, **၆၀**% , etc.
DIGIT	digit by digit	**+၉၅-၉ ၄၃၀၅၄၄၈၅**, အမှတ် **၄၄၇**, etc.
RANGE	range	**၇၀ - ၈၀** ဒီဂရီဖာရင်ဟိုက်, ဒေါ်လာ **၁၀၀** နှင့် **၁၅၀** ကြား, etc.
SCORE	score	**၂-၃** ဂိုး, **၂း၃** ဂိုး, etc.
DIMENSION	dimension	ပေ **၄၀ × ၆၀**, etc.
NRC	national identification number	**၅**/မရန(နိုင်) **၁၂၃၄၅၆**, etc.

III. Defining Rules for Number Normalization

For detection of the semiotic classes defined in Table I, 60 rules are defined for simple rule-based number normalization system. Rule-based number normalization is implemented by regular expression (RE) in Perl language. Look-up dictionary is used for expansion of measurement units and currency unit symbols. An example RE for detecting date range in DATE semiotic class is as follows:

if($inputdata =~ m/(\s+[၀-၉]{4}\s*)-(\s*[၀-၉]{4}\s*)($strTwoDateSuf)/g)

RE for detecting currency with currency suffix is written as follows:

if($inputdata =~ m/([၀-၉]{1,3},([၀-၉]{3},)*[၀-၉]{3})(\.?[၀-၉]*\s*)($strCurrencySuf)/g)

This rule-based number normalization system is used as a baseline in this paper.

IV. Weighted Finite State Transducer for Number Normalization

The main point of this paper is language-specific grammar that is compiled to WFST. A WFST consists of a set of states and transitions between states. Each transition is labeled with an input symbol from an input alphabet; an output symbol from an output alphabet; an origin state; a destination state; and a weight [13]. Finite State Transducer (FST) can represent certain sets and binary relations over string. A Deterministic Finite Automaton recognizes a regular language and a FST recognizes a regular relation. FSTs have been used to model the morphology, phonology and other text-analysis operations such as numeral expansion. WFSTs have been used in Text analysis model for Text-to-Speech (TTS) of the multilingual Bell Labs TTS system [14]. Kestrel text normalization system, a component of the Google text-to-speech synthesis system, also used grammars based on WFSTs for text normalization of many languages [3].

OpenGrm Thrax Grammar Compiler[2] [13] is used for the development of Myanmar number normalization system. The Thrax grammar compiler compiles grammars that consist of regular expressions, and context-dependent rewrite rules, into FST archives of weighted finite state transducers. It uses OpenFst library to provide an efficient encoding of grammars and general algorithms.

V. Myanmar Number Normalization

Myanmar number normalization is achieved by implementing two phases : classification and verbalization phases. Both phases are accomplished by defining grammars that are compiled to WFST. Semiotic classes defined in Section II is used for classification. Number prefixes and suffixes are also used as the clues for identifying semiotic class of the current word. Word segmentation process is needed as the pre-processing step of this system. Detailed processes for the whole system are discussed in the following sections.

[1] available at http://www2.nict.go.jp/astrec-att/member/mutiyama/ALT/index.html

[2] available at http://www.openfst.org/twiki/bin/view/GRM/Thrax

A. Grammars for Classification

Grammars that are compiled to the libraries of WFST are applied for classification phase. Before applying these grammars, some pre-processing rules are written for combining tokens. Some digit sequences are need to combined for classifying their semiotic classes. For example, in Myanmar sentence,

"မီတာ **၁၀၀ ၊ ၂၀၀ ၊ ၃၀၀** စီ ဝေးပါတယ်။"

(Each fars 100, 200, 300 meters.)

In this case, "**မီတာ**" (meter) is the clue for identifying the whole digit sequence as the NUMBER class and it is located before the first digit. The other digits need to be related that clue. Spaces between these digit sequences are removed as the pre-processing step and it becomes "မီတာ ၁၀၀၊၂၀၀၊၃၀၀ စီ ဝေးပါတယ်။". Therefore, the whole digit sequence can be identified as the NUMBER class.

For DATE semiotic class, more than one rule is needed to implement because many patterns of DATE are commonly written in Myanmar. Table II shows some examples of Myanmar date.

Table II
Myanmar Dates

Myanmar	English
ဇန်နဝါရီ **၄**၊ **၂၀၁၇**	January 4 2017
၂၀၁၇ ၊ ဇန်နဝါရီလ **၄**	2017 January 4
၂၀၁၇ ခုနှစ် ၊ ဇန်နဝါရီလ **၄**	2017 January 4
၂၀၁၇ ၊ **၄** လပိုင်း	4/2017
ဇန်နဝါရီလ **၄**	January 4
၄.၁.၂၀၁၇	4.1.2017
၄-၁-၂၀၁၇	4-1-2017
၄/၁/၂၀၁၇	4/1/2017
၁၃၇၈ ခုနှစ် ဝါဆို လဆန်း **၁၀** ရက်	10, early Warso, 1378
၁၃၇၈ ခုနှစ်၊ တန်ဆောင်မုန်း လပြည့်ကျော် **၁၀** ရက်	10, late Tazaungmone, 1378
၂၀၁၆-၂၀၁၇ ခုနှစ်	2016-2017
၂၀၁၆-၂၀၁၇ ပညာသင်နှစ်	2016-2017 academic year

There are two types of clues in classification for CURRENCY semiotic class. It can be used currency symbol as the prefix of the number or currency text as the prefix or suffix of the number. As an example, "$ **၅၀**","ဒေါ်လာ **၅၀**" and "**၅၀** ဒေါ်လာ" has the same meaning in English as "$50".

In this paper, cardinal, decimal and measure are defined in the member of NUMBER semiotic class. In Myanmar, ordinal numbers are usually written by using ordinary number or text.

Examples are "**၂** ကြိမ်မြောက်" (second) and "**ပထမ**အကြိမ်" (first).

The first example is ordinal number using ordinary number that means "second" in English and the second one is using text that means "first" in English. Therefore, it is no need to identify as one type of semiotic class. Measure unit symbols, 96 number prefixes and 275 number suffixes are used for classification of NUMBER semiotic class. These prefixes and suffixes are extracted by manual collecting of Myanmar text corpus. Measure unit symbols with their expansions like "**%** (ရာခိုင်နှုန်း)", "**°F** (ဒီဂရီဖာရင်ဟိုက်)", "**ft** (ပေ)" are applied by using StringFile function.

Finally, grammars for all defined semiotic classes are compiled to WFSTs and then these WFSTs are exported and used in Classification phase. Weights are assigned for defining different priorities of the classification. Example input string and output string of classification phase are as follows:

Input String: **၂၀၁၇** ဧပြီလ **၁၇** ရက်သည် မြန်မာ နှစ်ဆန်းတစ်ရက်နေ့ ဖြစ်ပါသည်။

(17th April, 2017 is Myanmar New Year's day.)

Output String : <DATE>year: **၂၀၁၇** month: ဧပြီလ day: **၁၇**</DATE> ရက်သည် မြန်မာ နှစ်ဆန်းတစ်ရက်နေ့ ဖြစ်ပါသည် ။

B. Grammars for Verbalization

Verbalization is the expansion of tokens into standard words and its expansion mainly depends on its semiotic class. Verbalization of semiotic classes that come out from classification phase is also done by using WFST. In Myanmar language, some symbols have to be neglected in verbalization. For example, in DIMENSION semiotic class,

"ပေ **၄၀** × **၆၀**" (40 ft × 60 ft) can be expanded into
"ပေ လေး ဆယ် ခြောက် ဆယ်".

In this case, "×" symbol must be neglected in verbalization.

A goal score, "**၃-၁** ဂိုး" (3:1) can be expanded into
"သုံး ဂိုး တစ် ဂိုး".

In the above SCORE case, "**-**" symbol needs to be replaced by suffix "ဂိုး". In Myanmar, "**-**" is usually used in SCORE case.

Myanmar people sometimes omit "တစ်" (one) in pronunciation of Myanmar digit sequence like "တစ်ဆယ့်" (ten) and "တစ်ထောင့်" (one thousand). Examples are:

၁၁း၀၀ နာရီ (11:00 hour)	-	ဆယ့် တစ် နာရီ
၁၉၁၅ (1915)	-	ထောင့် ကိုး ရာ့ ဆယ့် ငါး
၁၅၀၀ ကျပ် (1500 kyats)	-	ထောင့် ငါး ရာ ကျပ်

Therefore, some rules for fixing these particular cases are written in verbalization grammars.

C. Myanmar Number Names Expansion

All the grammars except digit by digit expansion depend on the expansion of Myanmar number names. The number name grammars depend on the factorization of digit string into sum of products of powers of ten.

These factorizations can be done by using Thrax Factorizer grammar. Number verbalization of these factorized strings into number names depends on the language. Most Western languages have no terms for 10^4, they say ten thousand. The factorization becomes $1\times10^1\times10^3$. In Myanmar language, we have a term "သောင်း" (ten thousand) for 10^4 and the factorization is 1×10^4. Although there are terms "သန်း" (million) for 10^6 and "ကုဋေ" (ten million) for 10^7 in Myanmar language, it is not commonly used in pronunciation of CURRENCY class. Myanmar people usually say "၁၀ သိန်း" (ten lakh) for 10^6 and "သိန်း ၁၀၀" (hundred lakh) for 10^7. Therefore, factorization defines $1\times10^1\times10^5$ for the first case and $1\times10^2\times10^5$ for the second case and Myanmar number names grammar has to be handled these factorizations into appropriate Myanmar standard words. As an example,

"**၁,၂၃၄,၅၆၇,၈၉၀**" (1,234,567,890) will be expanded into standard words as

"တစ် သောင်း နှစ် ထောင့် သုံး ရာ့ လေး ဆယ့် ငါး သိန်း ခြောက် သောင်း ခုနစ် ထောင့် ရှစ် ရာ့ ကိုး ဆယ်".

(twelve thousand three hundred forty-five lakh sixty-seven thousand eight hundred and ninety)

D. Finalization

After verbalization phase, some digit sequences are missing to identify their semiotic classes and to verbalize their standard words because they have no clue for defining semiotic classes. In finalization, post-processing have been done for fixing these cases. Some assumptions are made to decide whether current token should be classified as cardinal number or as digit by digit. If there is a comma or dot in digit sequence or it has one non-zero digit followed by many zero digits, it will be expanded into cardinal number names. If not, how many digits in the sequence will be checked. If it has three or more digits, it will be pronounced as digit by digit and if it has less than three digits, pronounced as the cardinal number. Myanmar people usually pronounce as that way. For example, Bus no. "**၃၉**" (39) is usually pronounced as "သုံး ဆယ့် ကိုး" (thirty-nine) and Bus no. "**၁၂၄**" (124) as "တစ် နှစ် လေး" (one two four).

VI. Sample Grammar Fragments

In classification phase, as we mentioned in Section V-A, some clues are used for classifying semiotic class. However, if there is a number sequence of same semiotic class, it is common that clue is located at the start or end of the sequence of numbers separated by Myanmar punctuation "၊". As an example, a sequence of currency number is written as

"**၂၀,၀၀၀ ၊ ၁၅,၀၀၀ ၊ ၁၀,၀၀၀** ကျပ်" (20,000 , 15,000 , 10,000 kyats) instead of writing "**၂၀,၀၀၀** ကျပ်၊ **၁၅,၀၀၀** ကျပ်၊ **၁၀,၀၀၀** ကျပ်" (20,000 kyats, 15,000 kyats, 10,000 kyats).

Example grammar for detecting sequence of currency is as follows:

For example : ၂၀,၀၀၀ ၊ ၁၅,၀၀၀ ၊ ၁၀,၀၀၀ ကျပ် => <CURRENCY>integer: ၂၀,၀၀၀ ၊ integer: ၁၅,၀၀၀ ၊ integer: ၁၀,၀၀၀ currencySuffix: ကျပ် </CURRENCY>

```
currency_sequence =
   u.Ins["<CURRENCY>"]
   u.Ins[" integer: "]
   validnum
   (u.D["."]
   u.Ins[" fractional_part: "]
   d+)? "၊"
   (u.Ins[" integer: "]
   validnum
   (u.Del["."]
   u.Ins[" fractional_part: "]
   d+)? "၊" )*
   u.Ins[" integer: "]
   validnum
   (u.Del["."]
   u.Ins[" fractional_part: "]
   d+)?
   u.Del[" "*]
   u.Ins[" currencySuffix: "]
   currency_suffix
   u.Ins["</CURRENCY>"]
```

In verbalization, Myanmar number name grammar is used as the basic and some rewrite rules are written to fix some particular cases. For example, omitting the pronunciation of "တစ်" before "ဆယ့်" and "ထောင့်" is written by using content-dependent rewrite rule and composition is applied in this case. Example code fragment is as follows:

For example : ၁၁-၃-၁၉၁၇ (11-3-1917) => ဆယ့် တစ် ရက် သုံး လ ထောင့် ကိုး ရာ့ ဆယ့် ခုနစ်

```
day = n.MYANMAR_NUMBER_NAME;
month_num = n.MYANMAR_NUMBER_NAME;
year = n.MYANMAR_NUMBER_NAME;
date_num =
   u.Del["<DATE>"]
   u.Del[" day: "]
   day
   u.Ins[" ရက် " ]
   u.Del[" month: "]
   month_num
   u.Ins[" လ "]
   u.Del[" year: "]
   year
   u.Del["</DATE>"];
remove_tit = CDRewrite["တစ်" : "" , "", (" ဆယ့် ") | (" ထောင့် "), sigma_star];
export DATE_VERBALIZE = Optimize[date_num @ remove_tit];
```

VII. Test Data Preparation

Two test data TestData-1 and TestData-2 are prepared for evaluating current WFST-based Myanmar number normalization. TestData-1 has 1,000 sentences which are randomly selected from ALT corpus that need to be normalized. For TestData-2, some Myanmar sentences from the Web that contains 947 sentences with Myanmar digits are selected. For these two test data, parallel tagged corpus is prepared for evaluating classification and parallel normalized corpus for evaluating the overall performance.

VIII. Experimental Results

Experiments are designed to test the performance of classification and verbalization of Myanmar number normalization.

A. Classification

The following formula is used for evaluating the performance of classification.

$$tag\ accuracy(\%) = \frac{\begin{array}{c}number\ of\ particular\ tags\\ in\ test\ data\end{array}}{\begin{array}{c}number\ of\ particular\ tags\\ in\ reference\ data\end{array}} \qquad (1)$$

Table III shows the number of particular tags in test data and reference data, and the percentage of tag accuracy. The overall tag accuracy is **94.3%** on TestData-1 and **92.6%** on TestData-2.

Table III
Classification Accuracy of Two Test Data

Semiotic Class	TestData-1	TestData-2
NUMBER	755/810 (93.2%)	725/789 (91.9%)
DATE	396/404 (98%)	422/442 (95.5%)
TIME	78/80 (97.5%)	68/68 (100%)
CURRENCY	71/82 (86.6%)	98/118 (83.1%)
DIGIT	5/5 (100%)	3/3 (100%)
RANGE	2/2 (100%)	6/8 (75%)
SCORE	5/8 (62.5%)	-
DIMENSION	-	2/2 (100%)
Overall Accuracy	1312/1391 (94.3%)	1324/1430 (92.6%)

B. Verbalization

Verbalization results is reported in terms of word error rate (WER), a standard for evaluation of Automatic Speech Recognition system. Simple rule-based number normalizer developed by Perl language is used as the baseline in this paper. In Table IV, current number normalization system based on WFST achieves WER **0.5%** for TestData-1 and **1.4%** for TestData-2, and which is **5.0%** and **6.3%** lower than the baseline system respectively.

Table IV
Overall Performance

	TestData-1 (WER%)	TestData-2 (WER%)
Baseline	5.5%	7.7%
WFST-based	0.5%	1.4%

IX. Error Analysis and Discussion

Some errors are found in classification phase in our experiments because of two factors. The first factor is that there is no clue before and after the digit sequence.

In this example, **၇၀** ကနေ **၃၁၀** ကီလိုမီတာ (from 70 to 310 kilometers) should be classified as

<NUMBER>**၇၀**</NUMBER> ကနေ <NUMBER> integer: **၃၀၀** numberSuffix: ကီလိုမီတာ</NUMBER>.

However the classifier cannot classify which class should be occupied "**၇၀**" because it has no clue.

The second factor is that there are some common prefixes or suffixes on both CURRENCY and NUMBER classes. For example, in this Myanmar sentence,

"ကမ္ဘာ့လူဦးရေက လတိုင်း **၆** သန်း လောက်တိုးပွားနေပါတယ်။" (The world's population has increased about 6 million per month.),

although "**၆** သန်း" (6 million) should be classified as the NUMBER class, its result class is CURRENCY because "သန်း" (million) is also the prefix of CURRENCY class.

As the same way, CURRENCY class is sometimes misclassified as NUMBER class. For example, in the sentence,

"အမေရိကန်ဒေါ်လာ သန်းပေါင်း **၃၈၀** ခန့် ရင်းနှီးမြှုပ်နှံမည်" (They will be invested approximately 380 million dollars.),

the prefix of the digit sequence "သန်းပေါင်း" is the prefix of CURRENCY class and the suffix "ခန့်" is the suffix used in NUMBER class. Therefore, classifier based on FST chose the shorter path "**၃၈၀** ခန့်" as the NUMBER class.

Although there are some classification errors, verbalization results of some misclassification have also correct pronunciation because almost same pronunciation for CURRENCY and NUMBER class in Myanmar language and post-processing of our number normalization. Therefore, low WER on overall accuracy is achieved in our experiments.

X. Conclusion and Future Work

Number normalization is a very important module in Text-to-Speech system. In this paper, semiotic classes for Myanmar language are identified and WFST-based Myanmar number normalization designed for Myanmar TTS system is implemented for the first time. Experimental results show that this WFST-based approach can get acceptable results for the performance of Myanmar number normalization and this can be used practically by integrating in Myanmar TTS system.

For solving misclassification and missing tag cases in classification phase, this phase will be replaced by applying sequence-to-sequence modeling. A tagged corpora for various types of data resources will be built for that modelling.

References

[1] P. Taylor, *Text-to-speech synthesis.* Cambridge university press, 2009.

[2] R. Sproat, A. W. Black, S. Chen, S. Kumar, M. Ostendorf, and C. Richards, ``Normalization of non-standard words,'' *Computer speech & language*, vol. 15, no. 3, pp. 287--333, 2001.

[3] P. Ebden and R. Sproat, ``The kestrel tts text normalization system,'' *Natural Language Engineering*, vol. 21, no. 03, pp. 333--353, 2015.

[4] Y. K. Thu, W. P. Pa, J. Ni, Y. Shiga, A. Finch, C. Hori, H. Kawai, and E. Sumita, ``Hmm based myanmar text to speech system,'' in *Sixteenth Annual Conference of the International Speech Communication Association*, 2015.

[5] S. Beliga and S. Martinčić-Ipšić, ``Text normalization for croatian speech synthesis,'' in *MIPRO, 2011 Proceedings of the 34th International Convention.* IEEE, 2011, pp. 1664--1669.

[6] F. Alam, S. Habib, and M. Khan, ``Text normalization system for bangla,'' BRAC University, Tech. Rep., 2008.

[7] T. Zhou, Y. Dong, D. Huang, W. Liu, and H. Wang, ``A three-stage text normalization strategy for mandarin text-to-speech systems,'' in *Chinese Spoken Language Processing, 2008. ISCSLP'08. 6th International Symposium on.* IEEE, 2008, pp. 1--4.

[8] K. Panchapagesan, P. P. Talukdar, N. S. Krishna, K. Bali, and A. Ramakrishnan, ``Hindi text normalization,'' in *Fifth International Conference on Knowledge Based Computer Systems (KBCS).* Citeseer, 2004, pp. 19--22.

[9] R. Sproat, ``Lightly supervised learning of text normalization: Russian number names,'' in *Spoken Language Technology Workshop (SLT), 2010 IEEE.* IEEE, 2010, pp. 436--441.

[10] T.-T. T. Nguyen, T. T. Pham, and D.-D. Tran, ``A method for vietnamese text normalization to improve the quality of speech synthesis,'' in *Proceedings of the 2010 Symposium on Information and Communication Technology.* ACM, 2010, pp. 78--85.

[11] R. Sproat and N. Jaitly, ``Rnn approaches to text normalization: A challenge,'' *arXiv preprint arXiv:1611.00068*, 2016.

[12] H. Riza, M. Purwoadi, Gunarso, T. Uliniansyah *et al.*, ``Introduction of the asian language treebank.'' Oriental COCOSDA, 2016.

[13] B. Roark, R. Sproat, C. Allauzen, M. Riley, J. Sorensen, and T. Tai, ``The opengrm open-source finite-state grammar software libraries,'' in *Proceedings of the ACL 2012 System Demonstrations.* Association for Computational Linguistics, 2012, pp. 61--66.

[14] R. Sproat, ``Multilingual text analysis for text-to-speech synthesis,'' in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, vol. 3. IEEE, 1996, pp. 1365--1368.