Acoust. Sci. & Tech. **32**, 5 (2011)

PAPER

Myanmar text-to-speech system with rule-based tone synthesis

Kyawt Yin Win and Tomio Takara*

Department of Information Engineering, University of the Ryukyus, 1 Senbaru, Okinawa, 903–0213 Japan

(*Received 5 August 2010, Accepted for publication 4 February 2011*)

Abstract: We have introduced a novel Myanmar text to speech (MyanmarTTS) system with rule-based tone synthesis. Myanmar is a tonal language that possesses unique characteristics compared with other tonal languages such as Chinese, Vietnamese and Thai. Such languages have complicated fundamental-frequency (F_0) patterns of tones, and F_0 is of foremost importance. Myanmar tones are unique in their simplistic pattern related not only to F_0 but also, more specifically to duration. Myanmar tones have different durations between short-tone and long-tone groups. In accordance, we defined a tone rule employing two parameters F_0 at the center of the syllable and the syllable's duration. The rule is implemented with a linear F_0 pattern. Large variability exists in the F_0 and duration uttered by different speakers of different syllables. Hence, for tone synthesis, normalization of the F_0 and duration is important and necessary to discriminate tones. We proposed a normalization method and the effectiveness of this method was confirmed in the distribution of the F_0 and duration. The intelligibility of the synthesized tone was confirmed through listening tests with correct rates of 95.6% for male and 97.8% for female speech. As a result, we showed that the linear pattern is sufficient for Myanmar tone synthesis.

Keywords: Myanmar, Tonal languages, Text-to-speech, Tone synthesis, Normalization, Rule-based

PACS number: 43.72.Ja [doi:10.1250/ast.32.174]

1. INTRODUCTION

There has been ample research on speech processing in some tonal languages such as Chinese, Thai and Vietnamese. Nowadays, many applications in various information systems incorporate digital speech technologies, such as telecommunications, security and warning systems, telephone inquiry systems, multimedia applications, electronic mail reading, and aids for the blind and people with speech impairments. This type of communication is progressing [1]. For Myanmar language, keeping track of world trends in this field is important for fostering cooperation in the field of global technology. For this purpose, the Myanmar text-to-speech (TTS) system is essential. While TTS systems for major world languages are quite advanced, there has not been any MyanmarTTS system and research on speech synthesis is lacking.

In this paper, the first MyanmarTTS system is introduced. The MyanmarTTS system is a rule-based speech synthesis system, in which fundamental speech units are demi-syllables with a Level tone that has a flat fundamental frequency pattern. We use the Myanmar (Burmese) official language for the TTS system. To construct the TTS system, monosyllabic words are analyzed and the parameters are obtained for the synthesis of Myanmar tones. This TTS system uses a source filter model and furthermore, a log magnitude approximation Filter [2] is employed to reproduce the vocal tract character. Myanmar is a tonal language comprising 4 different tones. The Level and High falling tones belong to the long-tone group and the Falling and Checked tones belong to the short-tone group. The Checked tone is accompanied by a glottal stop like the Broken tone and the Drop tone of Vietnamese. Standard Myanmar is used by 8 main races and subraces as their official language. It is spoken in most of the country with slight regional variations. Accordingly a large variability exists in the F_0 and durations among the speakers.

In Myanmar, however, tones are unique in their simplistic pattern related to not only F_0 but also, more specifically and importantly, to duration. This is the basis for the proposed linear pattern for the tone rule using the normalized F_0 and duration of each tone among the utterances. In this study, normalization method is different

*e-mail: takara@ie.u-ryukyu.ac.jp

from those of other TTS systems such as for Thai and Vietnamese [3,4]. The synthesized tones are evaluated by employing listening tests. The results show that our rule-based linear pattern gives high intelligibility of the synthetic sound and is considered to be sufficient for the MyanmarTTS system.

2. MYANMAR LANGUAGE OVERVIEW

Myanmar belongs to the Lolo-Burmese sub-branch of the Tibeto-Burmese branch of the Sino-Tibetan language family. Myanmar script draws its source from Brahmi script, which flourished in India from about 500 B.C. to over 300 AD. Myanmar is a tonal language. This means that all syllables in Myanmar have prosodic features that are an integral part of their pronunciation. Standard Myanmar is based on the dialect spoken in the lower valleys of the Irrawaddy and Chindwin rivers. It is spoken in most of the country with slight regional variations. In addition, there are other regional variants that differ from standard Myanmar in pronunciation and vocabulary. Furthermore, there are two registers: formal and colloquial. The formal register is used in official publications, radio and TV broadcasts, literary works, and formal speech. The colloquial register is used in daily communications.
In Myanmar there are 8 main races and 135 subraces. Myanmar (Burmese) is the official language in Myanmar. We choose the formal register for speech synthesis.

Myanmar has 4 tones and a simple syllable structure that consists of an initial consonant followed by a vowel with an associated tone. Different tones denote different meanings for syllables with the same structure of phonemes. In the Myanmar writing system, a tone is presented by a diacritic mark. The four Myanmar tones are shown in Table 1. Figure 1 shows an example of the F_0 contour of the four Myanmar tones with the syllable /ma/ uttered by a male native speaker. More details on the Myanmar language can be found in [5–7]. There are 33 consonants and three groups of vowels in Myanmar pure vowel, nasal vowel and diphthong. There are 8 monophthongs and 4 diphthongs in basic vowels of Myanmar. The Myanmar syllable has the structure shown in Fig. 2. The syllable has the phonemic shape of C (G) V (N/ʔ) T, where an initial consonant C is obligatory, a glide consonant G is optional, a vowel V is obligatory, a final consonant-nasal N or stopped ʔ is optional, and tone T is obligatory, respectively. The minimum syllable is CVT. Syllables in the Myanmar language consists of either a vowel by itself or a consonant combined with a vowel.

3. MYANMAR TTS SYSTEM

The design of the MyanmarTTS system is shown in Fig. 3. This design is based on a general speech synthesis system [8]. The input text comprises Myanmar-script-equivalent characters. The output is synthetic sound. The proposed system can synthesize a syllable and a word and a sentence.

In our system, the fundamental speech consisted of demi-syllables that were acquired by dividing a syllable with a cut-off point at 100 ms of the vowel part from the boundary of C and V. There are about 800 demi-syllables in Myanmar. For the purpose of compiling the database for

Table 1 Four Myanmar tones.

Tone name	Symbol	Description
Level	◌	/là/ - moon
Falling (Creaky)	◌	/la/ - come
High Falling (High)	◌း	/lá/ dray horse
High Extremely short (Checked)	◌်	/laʔ/ - fresh

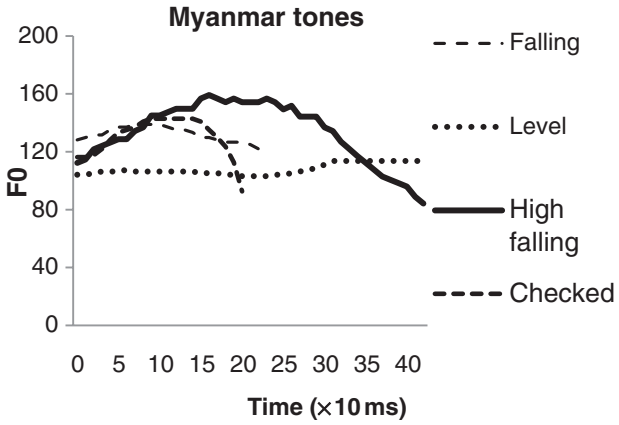

Fig. 1 Examples of four tones of Myanmar syllable /ma/ uttered by a male native speaker.

Initial consonant	Glide consonant	vowel	Final consonant	Tone
C	(G)	V	(N/ʔ)	T

Fig. 2 Myanmar syllable structure.

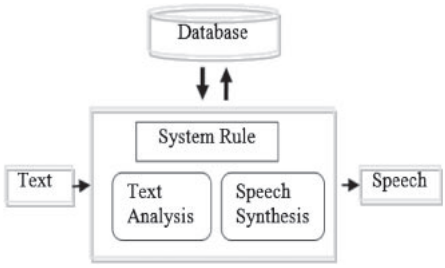

Fig. 3 Diagram of Myanmar TTS system.

Table 2 Examples of Myanmar phonemes and consonantal letters used in the system.

Phonemes	Myanmar letters	Character input
ə	အ	a
k	က	ka
k^h	ခ	kha
g	ဂ	ga
ŋ	င	nga
s	စ	sa
s^h	ဆ	sha
z	ဇ	za
ɲ	ည	nya
t	တ	ta
t^h	ထ	hta
d	ဒ။ဓ	da
n	နာ။ဏာ	na
p	ပ	pa
p^h	ဖ	hpa
b	ဗ။ဘ	ba
m	မ	ma
j	ယ	ya
ɹ	ရ	ra
l	လ	la
w	ဝ	wa
θ	သ	tha
h	ဟ	ha
ʦ	ကျ ။ ကြ	ca
ʧ	ချ ။ ခြ	cha
ʤ	ဂျ ။ ဂြ	gya
ɬ	လှ	hla

(/a/ is used as an example of syllables)

the speech unit, Myanmar demi-syllables were collected and their sounds were prepared by recording on a digital audio tape (DAT) at a 48 kHz sampling rate with a 16-bit resolution. After that, they were down-sampled to 10 kHz for analysis.

3.1. Text Analysis

The purpose of our text analysis was to extract the phonetic and prosodic features. We developed Myanmar pronunciation lists (database) using Myanmar-script-equivalent English alphabets. The input text is Myanmar-script-equivalent English alphabets, as shown in Table 2. With the rule for mapping, the sequence of mapped characters is obtained from the list and then the information of the character is retrieved. We used English alphabets as the target of mapped sequences and made a mapping table for 4 tones, as shown in Table 3.

Table 3 Rule for mapping in text analysis.

Marks	Number
Level tone	0
Falling tone	1
High falling tone	2
Checked tone	3

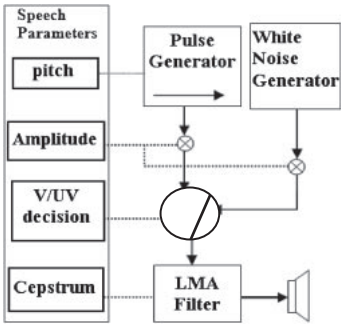

Fig. 4 MyanmarTTS's speech synthesis subsystem.

For example, the sentence "သူမလှသည်" which means "She is beautiful," is mapped into a sequence of syllables as "thu0ma1hla1thi0."

3.2. Speech Analysis and Synthesis

3.2.1. Speech analysis

MyanmarTTS adopts a short-time cepstral analysis. The frame duration is 25.6 ms and the frame interval or frame shifting time is 10 ms. A time–domain Hamming window with a length of 25.6 ms was used.

The cepstrum is defined as the inverse Fourier transform of the short-time logarithmic spectrum [9]. Cepstral analysis is advantageous as it can separate the spectral envelope part and the excitation part. The resulting parameter of speech units include the number of frames and, for each frame, voiced/unvoiced (V/UV) decision, pitch period and cepstral coefficients $c(m)$, $0 \leq m \leq 29$.

3.2.2. Speech synthesis

Under the control of the synthesis rule, the speech synthesis subsystem generates speech from prestored parameters. The source-filter model [10] is used as the speech production model. Figure 4 shows the structure of the speech synthesis subsystem in MyanmarTTS. The synthetic sound is produced using the log magnitude approximation (LMA) filter, which was introduced by Imai [2]. It presents the vocal tract characteristics. The spectral envelope is represented by the cepstral coefficients of 30 lower-order elements. The LMA filter is a pole-zero filter that is able to represent efficiently the vocal tract features for all speech sounds.

We synthesized speech from prestored parameters with over 800 demi-syllables as speech units under the system rules. After synthesis, the waveform was displayed and played.

3.3. Intelligibility Test

In order to evaluate the intelligibility of the MyanmarTTS system, a syllable intelligibility test was carried out. To create a set of data, we referred to the JEIDA guideline (JEIDA-G-24-2000). The test methodology is described by Imai and Abe [11].

3.3.1. Speech materials

The test consisted of 248 monosyllable words. The test stimuli are analysis-synthesis sounds with level tone which has a flat F_0 pattern. Meaningful monosyllable words were mostly used as well as some meaningless monosyllable words were used. Those monosyllable words are analysis-synthesis sounds with level tone, disregarding other tones. The total number of sounds was 496 (i.e., 31 consonants $\times$ 8 vowels $\times$ 2 repetitions).

3.3.2. Method

Five native Myanmar listeners participated in listening tests. All listeners had normal hearing ability. In the test, monosyllable words were presented randomly. The listening tests were performed in a soundproof room and the listeners used headphones (SONY MBR-XB700). Instructions for the test were prepared and the listener performed the practice test once. Each listener performed the test twice. The listener listened word by word and wrote the word they heard on the answer sheet. Each word was played one time within a 2 second interval with a pause time of 3 seconds to write the answer. The time taken for each test is 48 minutes.

3.4. Result

The average intelligibility score were 92.56%. We found that most of the error came from nasal consonant words and dental consonant words. The results of the intelligibility test are as shown in Figs. 9 and 10. The score of the intelligibility test for Japanese [11] is 91.2%. Compared with Japanese intelligibility test, we consider our score to be reasonable and sufficient for syllable intelligibility.

The results of correct answers for each syllable are shown in Fig. 10. The scores of correct answers for each syllable were reasonably high except the score for /θ/ and /η/. The score for /θ/ was worse than those for other syllables, since /θ/ is very similar to /th/ and /t/, dental consonants. In addition, the meaningless syllables combined with /θ/ were used in the listening test, which causes confusion among these dental consonants.

Also, the score for /η/ is worse than those for other syllables, since /η/ is very similar to /mya/, a nasal consonant. Also meaningless syllables with /η/ were used in the listening test, which causes confusion among these nasal consonants. Moreover, /η/ is mostly used as the final consonant for VC and the ending consonant of sentences in the written language. In accordance with the above mentioned conditions, we conclude the intelligibility scores of /θ/ and /η/ was worse than those for other syllables.

4. TONE SYNTHESIS PROCEDURE

4.1. Tone Synthesis

The four tones were analyzed to extract F_0 patterns. The data set was prepared as voiced sounds and meaningful words. We selected consonant-vowel (CV) form with voiced consonants /b/, /m/, and /l/ and three typical vowels /a/, /i/, and /u/. The consonants /b/, /m/, and /l/ are voiced consonants which are in a different place of articulation and have a different manner of articulation. We think that they covered most kinds of voiced consonants. In total, 180 words (i.e., 3 consonants $\times$ 3 vowels $\times$ 4 tones $\times$ 5 speakers) were used for tone analysis. After analysis, four tones were distributed as shown in Fig. 5. We found that the four tone groups overlapped and were not clearly discriminated. Myanmar tones have simple F_0 patterns and different durations between short-tone and long-tone groups. The Falling and Checked tones belong to the short-tone group and high falling and level tones belongs to the long-tone group. The Checked tone has a very short duration and a glottal stop feature. Glottal stop in a speech synthesis system has been studied by Takara [12]. Consequently, we considered that the minimization of F_0 variations and duration variations are required for tone discrimination.

Myanmar tones are not only related to F_0 but also more specifically and importantly, to length. A change in duration constitutes a change in the feature of tone. For this reason, we normalized F_0 and duration to obtain relative values among the tones.

4.2. Normalization in F_0 and Length

We selected F_0 from three points at the center of the syllable for each tone and for each speaker. The average F_0 was defined for a male speaker, a female speaker and both male and female speakers to fix a standard value. Similarly, the average duration for each tone was defined. Subsequently, the F_0 and duration of each tone for each speaker was normalized to an approximated standard value. In this view, our normalization method is different and unique compared with other studies of tone synthesis research on, for example, Thai and Vietnamese [3,4]. One male speaker and two female speakers were Yangon (Capital of Myanmar) natives. One male speaker was a native of lower Myanmar and one female speaker was a native of upper Myanmar. Their average age was 35.

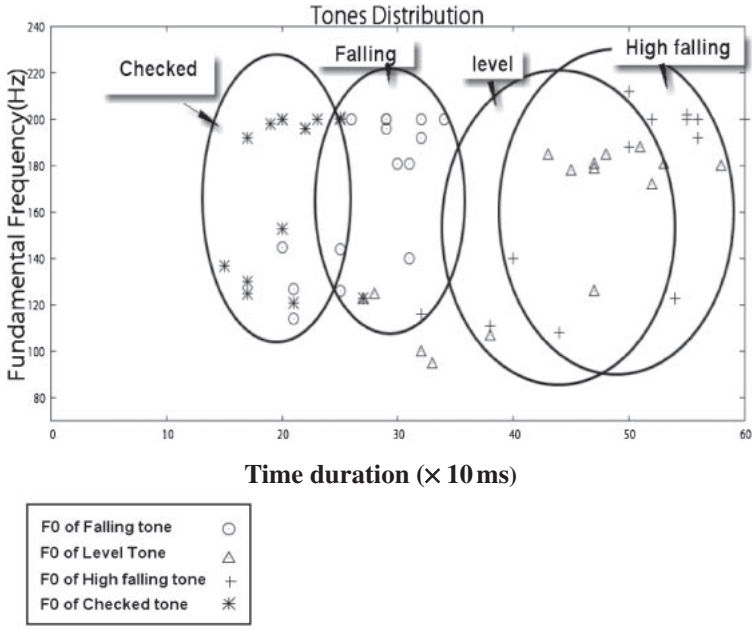

Fig. 5 Tone distributions of the analysis synthesis words "ba," "mi," and "lu" uttered by three female speakers and two male speakers.

To minimize large differences in F_0 and durations among the speakers by means of tones, normalization is carried out as follows. The average F_0 for each tone is given by

$$f_{ij} = \frac{1}{n}\sum_{k=1}^{n} f_{ij}^{(k)}. \quad (1)$$

where n is the number of F_0 points at the center of syllables. f_{ij} is the average F_0 from the center n points of syllable for ith tone and jth speaker. "k" is the time point extracted at the center of syllables. In this study, we set $n = 3$. The average F_0 of four tones is defined as A_j and the average of all speakers is defined as A_0.

$$A_j = \frac{1}{t}\sum_{i=1}^{t} f_{ij} \quad (2)$$

Here, t is the number of tones, which is four.

$$A_0 = \frac{1}{s}\sum_{j=1}^{s} A_j \quad (3)$$

Here, s is the number of speakers. In this research, $s = 5$ was used. We define the relative value R_{ij} as

$$R_{ij} = f_{ij} - A_j \quad (4)$$

${f_{ij}}^0$ is the ruled-based F_0 for the ith tone, and is defined as

$${f_{ij}}^0 = A_0 + R_{ij} \quad (5)$$

Similarly, the durations for each tone were normalized.

We plotted the results in Fig. 6 showing the distribution of four tones for both males and females after the normalization of F_0 and duration, which are clearly grouped and discriminated. These results mean that the duration and F_0 are equally important for discriminating the tones. From this figure, we confirmed that normalization is an effective method of discriminating the four tones. In Fig. 6, the normalized F_0 and duration of four tones for the words "ba," "mi," and "lu" are distributed. These distributions include the F_0 and duration for both males and females. To define general tone rules for male and female we calculated the average of two parameters, F_0 and duration, from the results in Fig. 6. The tone rules are implemented with F_0 linear patterns as shown in Fig. 7. We applied the Eqs. (4) and (5) in Fig. 7. The relative values of two parameters, F_0 and duration are used with some simplifications. We set the average F_0 value of the Level tone as F_b for the tone rule. F_b for males is 110 Hz and that for females is 160 Hz. We carried out listening tests to evaluate intelligibilities of tones for synthetic speech of syllables and to evaluate the effect of normalization using these rules.

5. EVALUATION BY LISTENING TEST

5.1. Listening Test of Tone Intelligibility

The intelligibility of tones was evaluated through three types of listening tests by three native listeners. We incorporated male and female speech.

Type 1: Rule-based sounds
- 1-1: Synthetic sounds: Rule-based male sounds
- 1-2: Synthetic sounds: Rule-based female sounds

Type 2: Analysis-Synthesis sounds

Type 3: Original sounds

All synthetic sounds used cepstra from speech units with Level tone. The word set comprises three typical vowels "a," "i" and "u" with voiced consonants "b," "m,"

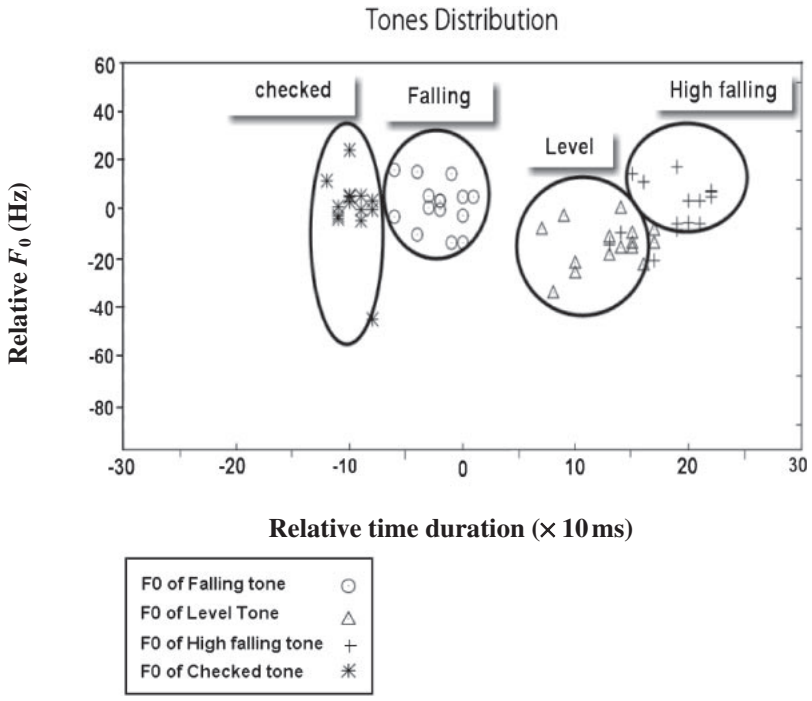

Fig. 6 Tone distributions of the words "ba," "mi," and "lu" uttered by three female speakers and two male speakers after normalization of F_0 and time (duration).

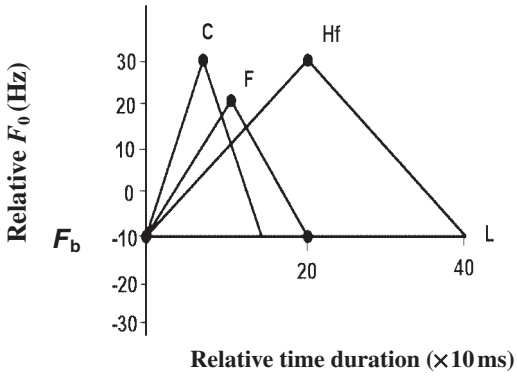

Fig. 7 Diagram of tone rule for males.
L: Level tone, F: Falling tone, Hf: High falling tone, C: Checked tone.

and "l," then syllable words were "ba," "mi," and "lu." Forty datums were prepared, given ten tokens for each tone and all sounds were meaningful words. The total number of sounds for each type was 240 (i.e., 3 words × 4 tones × 10 repetitions × 2 genders).

The listening test of Type 1, Type 2 and Type 3 were performed for each gender and for each word. In the all tests, each sound was played once at random and the listeners had to choose a word out of four possible words (tones) within 2 seconds. The listening tests were performed in a sound proof room. Among the three listeners, two were Yangon (Capital of Myanmar) natives and one was from the middle of Myanmar. They had normal hearing ability.

Table 4 Confusion matrices of tone synthesis.

Unit: %

Type 1 Rule-based		L	F	Hf	C
Male	L	96.8	0	2.6	0.6
	F	0.6	95.6	1.6	2.2
	Hf	8.5	0.5	91	0
	C	0.2	0	0	99.8
Female		L	F	Hf	C
	L	96.4	0	3.6	0
	F	0	97.5	0	2.5
	Hf	0.3	0	99.7	0
	C	0	0.3	0	99.7
Type 2 Analysis-Synthesis		L	F	Hf	C
	L	100	0	0	0
	F	0	95.6	0	4.4
	Hf	0	0	100	0
	C	0	0	0	100
Type 3 Original sounds		L	F	Hf	C
	L	100	0	0	0
	F	0	100	0	0
	Hf	0	0	100	0
	C	0	0	0	100

L: Level tone, F: Falling tone, Hf: High falling tone, C: Checked tone

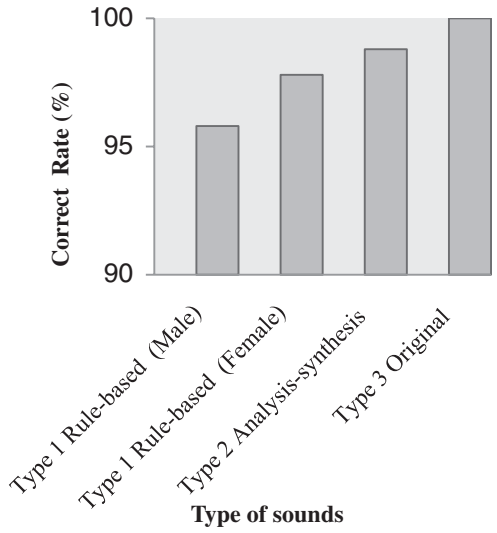

Fig. 8 Correct rates of perception of synthesized tone.

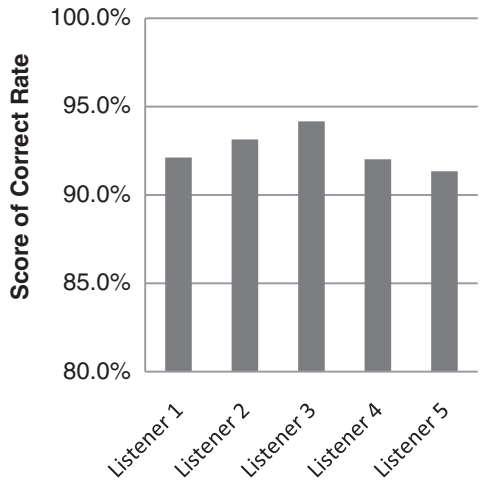

Fig. 9 Results of intelligibility test for Myanmar syllables.

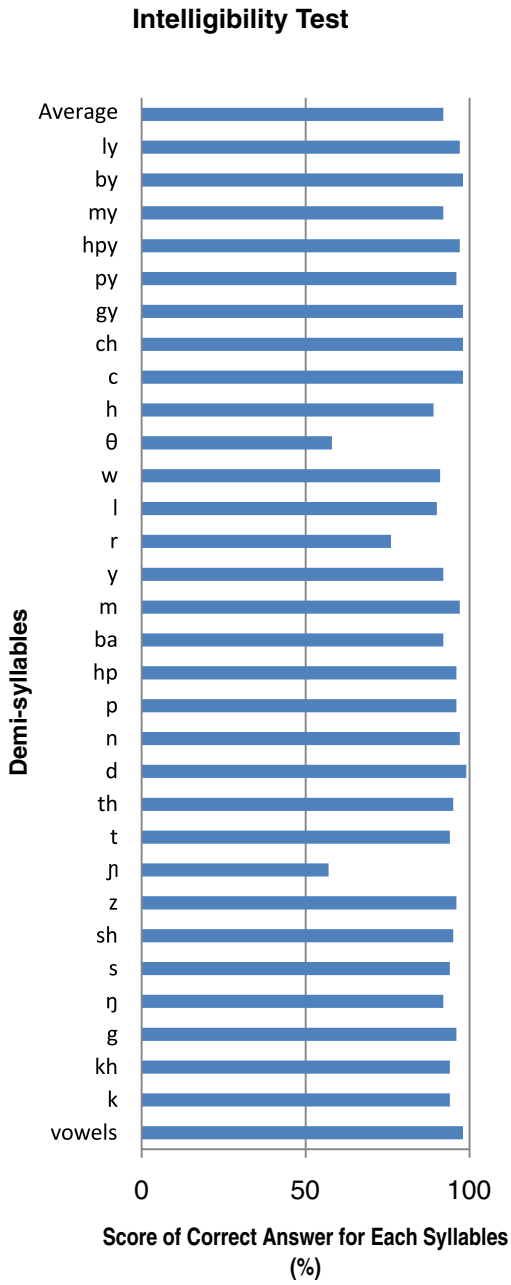

Fig. 10 Correct answers for each syllable.

6. RESULT AND DISCUSSION

Results of these tests are shown in Table 4 and Fig. 8. Table 4 shows confusion matrices of the tests. Figure 9 shows the correct rate of listening tests. The result of our tone synthesis system and the effectiveness of normalization are discussed as follows:

- Type 1, proposed method achieves a reasonably high correct rate of 95.8% for male speakers and 97.8% for female speakers. Type 2, analysis synthesis sounds, elicits the a high correct rate of 98.8% and Type 3, original sounds, elicits the highest correct rate 100%. These results indicate that the proposed method has a high intelligibility score for listening tests.
- The proposed method is similar to the F_0 linear pattern of the VieTTS system [4] with the same analysis-synthesis method. In the VieTTS system, the result for the linear pattern is about 85% for males, whereas the result of our system for males is 95.8%, even though our listening tests were done using the speech sounds of multiple speakers and different genders.
- Compared with those results, our linear pattern of tone rule is more effective than VieTTS's corresponding pattern because we adopted the normalization method for multiple speakers of different genders and for multiple syllables.
- From the confusion matrices of Type 1, we can see that errors occurred mostly between the High falling tone and the Level tone and between the Falling tone

and the Checked tone. In Type 1, we used the same length for the Level tone and the High falling tone. Furthermore, the spectra from the Level tone are used for all tones, whereas the original lengths of speech unit are used for Type 2 and Type 3. The results of confusion matrices show that the durations of tones are as important as F_0 for tone synthesis.

- Consequently, our proposed tone synthesis rule with a linear pattern is sufficient and effective for Myanmar tone synthesis for both male and female speech.

7. CONCLUSION

A novel Myanmar text-to-speech system with rule-based tone synthesis was introduced. This synthesis system incorporates short-time cepstral analysis. The proposed system can synthesize a syllable and a word and a sentence. We implemented tone rules of a linear pattern based on two parameters, the average value of F_0 at the center of the syllable and the syllable's duration. We confirmed that our proposed tone synthesis rule is sufficient and effective for Myanmar tone synthesis for the speech of both males and females. Myanmar tones are unique in their simplistic pattern related not only to F_0 but also, more specifically and importantly, to duration. In view of this, normalization by means of average F_0 and average length is useful for obtaining reliable values for tone discrimination and tone synthesis. We showed that duration is strongly related to tones, similarly to F_0, for Myanmar. The effectiveness of normalization is confirmed by distribution of tones and the intelligibility scores of listening tests. The introduced normalization method is applicable to other tone synthesis rules of other tonal languages.

There are, however, issues which still require investigation, such as the improvement of speech naturalness and the modification of the variations of F_0 and duration in continuous speech. An interesting theme that also remains is the adaptation of F_0 using spectra among speakers of different genders.

ACKNOWLEDGEMENT

The authors would like to express their deep gratitude to the referees for their valuable suggestions and their taken time for very convenient revisions. The authors thank to Myanmar native speakers who participated in the recording for speech data and in the listening tests in the evaluations.

REFERENCES

[1] T. Šef and M. Gams, "A complete Slovenian text-to-speech system," *Int. J. Speech Technol.*, **6**, 277–287 (2003).

[2] S. Imai, "Log magnitude approximation (LMA) filter," *Trans. IECE Jpn.*, **J63-A**, 886–893 (1980).

[3] P. Seresangtakul and T. Takara, "Analysis of pitch contour of Thai tone using Fujisaki's model," *Proc. ICASSP '02*, Vol. 1, pp. 505–5089 (2002).

[4] T. T. Do and T. Takara, "Vietnamese text-to-speech system with precise tone generation," *Acoust. Sci. & Tech.*, **25**, 347–353 (2004).

[5] Myanmar Language Committee, *Myanmar Grammar* (Myanmar Language Committee, Ministry of Education, Myanmar, 2005).

[6] U. Thein Tun, "Some acoustic properties of tones in Burmese," in *Papers in South-East Asian Linguistics 8: Tonation*, D. Bradley, Ed. (Australian National University, Canberra, 1982), pp. 77–116.

[7] J. K. Wheatley, "Burmese," in *The World's Major Languages*, B. Comrie, Ed. (Oxford University Press, New York, 1987), pp. 834–845.

[8] T. Takara and T. Kochi, "General speech synthesis system for Japanese Ryukyu dialect," *Proc. 7th Westpac*, pp. 173–176 (2000).

[9] A. M. Noll, "Cestrum pitch determination," *J. Acoust. Soc. Am.*, **41**, 293–309 (1967).

[10] S. Furui, *Digital Speech Processing, Synthesis, and Recognition*, 2nd ed. (Marcel Dekker, Inc., New York, 2001), pp. 30–31.

[11] S. Imai and Y. Abe, "Experimental speech synthesis from CV-syllable cepstral parameters," *Trans. IECE Jpn.*, **J64-D**, 861–868 (1981).

[12] T. Takara, "Experimental study on perception of the glottal explosive of the Japanese Ryukyu dialect," *Proc. EuroSpeech '95*, pp. 953–956 (1995).

Kyawt Yin Win received a B.Sc. degree in Physics from Malamyine University, Myanmar in 1992 and B.Ed. degree from Yangon University of Education, Myanmar in 1998. During 1992 to 2006 she has been employed as a senior teacher under the Ministry of Education, Myanmar. She received M.Ed. degree in Physics Education from Graduate school of Education, University of the Ryukyus, Japan in 2006. From 2007 to 2011 she studied in the Doctoral Course of the Department of Information Engineering, Graduate school of Engineering and Science, University of the Ryukyus, Japan. She is a student member of ASJ and ISCA. Currently she is a postdoctoral researcher at the Department of Information Engineering, University of the Ryukyus, Japan. Her research interests include speech processing and digital filters.

Tomio Takara received his B.S degree in Physics from Kagoshima University, Japan, in 1976 and M.E. degree and Dr. Eng degree in Information Processing from Tokyo Institute of Technology, Japan, in 1979 and 1983, respectively. During 1991–1992, he studied at Carnegie Mellon University as a visiting scientist. He has been a professor in the Department of information Engineering, University of the Ryukyus, Japan, since 1995. He received the 1990 Okinawa Society Award for Encouragement of Study on Okinawa. Currently he is the Dean of Faculty of Engineering, University of the Ryukyus. He is presently interested in spoken language processing and machine intelligence.