DEVELOPMENT OF THAI TEXT SET FOR TELEPHONOMETRY

Terdpong Daengsi¹, Apiruck Preechayasomboon², Gareth Clayton³, Chai Wutiwiwatchai⁴

¹Faculty of Information Technology, King Mongkut’s University of Technology North Bangkok, Bangkok, Thailand
²International Business Development Department, TOT Public Company Limited, Bangkok, Thailand
³Department of Applied Statistics, Faculty of Applied Science, King Mongkut’s University of Technology North, Bangkok, Thailand
⁴Human Language Technology Laboratory, National Electronics and computer Technology Center, Pathumthani Thailand

Email: therdpong1@yahoo.com, apiruck@tot.co.th, gareth@kmutnb.ac.th, chai.wutiwiwatchai@nectec.or.th

ABSTRACT

This paper presents the research to find the appropriate Thai text set to create a special Thai speech resource for telecommunication research, called telephonometry, using a survey. This will benefit the telecommunication in Thailand. The Thai text set was designed to be in compliance with the ITU-T standard. The data gathering was from the field survey results that had been done in 7 universities for general user representatives and 6 call centers for heavy user representatives, with 702 participants. 50 highest frequently used sentences (or phrases) were investigated using keywords for each possible context and then those were analyzed. Eventually, the final result can be applied as the Thai Text Set for Telephonometry (TTST) that will also be useful to create speech samples for quality measurement of VoIP services by telecom operators in Thailand, when VoIP services are used, and voice quality is important for users and becomes an issue.

Index Terms—speech resource; speech material; speech database; telephonometry; telecommunication

1. INTRODUCTION

There are some limitations in IP networks for transmitting VoIP which is a real-time application. Therefore, voice quality at the destination could be degraded. To find degraded voice quality, International Telecommunication Union – Telecommunication Standardization Sector (ITU-T) generally uses Mean Opinion Score (MOS) which is obtained from subjective and/or objective measurement methods. However, ITU-T has not focused on speech perception of native listeners who use tonal languages (e.g., Thai) that is different from speech perception of native listeners who use non-tonal languages (e.g., English) because changing tones can change words. For example, the Thai word “lín” (klaj) means “far” but the Thai word with another tone “línšt” (klajt) means “near”. Therefore, research with standard Thai spoken language should be conducted. However, it requires a special speech resource for the research areas of telecommunication directly.

This paper presents development of TTST which is a part of development of Thai Speech Set for Telephonometry (TSST). Firstly, some background information about the Thai brains, telecommunication, voice quality measurement and Thai language resources are presented in section 2. Then, methodology, results, discussion and conclusion are presented in section 3-6 respectively.

2. BACKGROUND

2.1. Language Ability and Tone Processing in Thai Brains

Language ability involves Broca’s and Wernicke’s areas in the cerebral cortex of Left Hemisphere (LH) which are responsible for speaking ability and language comprehension respectively [1-2]. A few papers have claimed that LH of Thai brains have an advantage when listening to Thai which is a tonal language. For example, the LH of brains of the participants who do not understand Chinese were predominant in perception of Thai speech samples which they understood, while Right Hemisphere (RH) of the brains of the participants were predominant in perception of Chinese speech samples [3]. Also, it was found that Thai native listeners presented activation in some parts of LH when compared with American English native listeners who do not understand Thai [4]. Whereas the results from another research, about presenting sample of Thai tones to groups of Thai native listeners and Chinese and English native listeners who do not understand Thai gave similar results [5].

2.2. Telecommunication Technology

2.2.1. 3G and 4G Technologies

Whereas, Thai people are trying to use 3G in Bangkok and some provinces, after official launching by TOT December last year, people in Scandinavia, Oslo and Stockholm, are starting to use 4G [6-7], as in figure 1. In the future, 4G technology will replace the earlier technologies with many advantages, such as, very high speed rate of data transfer and high capacity to support applications that require more resources [8-9]. It could be predicted that VoIP will be a basic service that is supported by 4G technology.
2.2. Updated Status of VoIP Technology in Thailand [11-16]
National Telecommunications Commission (NTC) acts as the regulator who provides the policy to support telecommunication in Thailand. NTC assigned 060-xxx-xxx for VoIP telephone numbering that will launch in the future already. Also, it has issued the QoS standard for VoIP to operators who will provide VoIP services. To share VoIP technology knowledge and educate telecom people, NTC provides training and offers scholarships to graduate students and researchers yearly. For operators, TOT started using VoIP technology behind the scene for Y-TEL 1234 for about 10 years ago. At present, TOT provides TOTnetcall and 008, whereas, True provides Nettalk and CAT009, using VoIP technology.

2.3. Voice Quality Measurement
ITU-T issued the voice quality measurement methods which could be mainly classified into subjective and objective measurement methods, as in figure 2. It is recommended to assess the voice quality by using MOS, the official scale that defines 5 for excellent, 4 for good, 3 for fair, 2 for poor and 1 for bad [17]. Voice quality of VoIP is affected by several network parameters, for example, delay, jitter, loss and echo. Besides, audio codec selection is also an important factor that affects VoIP Quality because a different audio codec offers a different MOS, as in table 1.

![Figure 2. Overview of Voice Quality Measurement Methods [17]](image)

<table>
<thead>
<tr>
<th>Codec</th>
<th>Algorithm</th>
<th>Bit-rate (Kbps)</th>
<th>Frame (ms)</th>
<th>Bits per frame</th>
<th>Codec Delay (ms)</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711</td>
<td>PCM</td>
<td>64</td>
<td>0.125</td>
<td>8</td>
<td>0.25</td>
<td>4.1</td>
</tr>
<tr>
<td>G.723.1</td>
<td>MP-MLQ</td>
<td>6.3</td>
<td>30</td>
<td>189</td>
<td>67.5</td>
<td>3.8</td>
</tr>
<tr>
<td>G.729</td>
<td>CS-ACELP</td>
<td>8</td>
<td>10</td>
<td>80</td>
<td>25</td>
<td>~3.9</td>
</tr>
<tr>
<td>G.729A</td>
<td>CS-ACELP</td>
<td>8</td>
<td>10</td>
<td>80</td>
<td>25</td>
<td>3.7</td>
</tr>
<tr>
<td>G.722</td>
<td>SB-ADPCM</td>
<td>48/56/64</td>
<td>0.0625</td>
<td>3-4</td>
<td>1.5625</td>
<td>~4.1</td>
</tr>
</tbody>
</table>

Table 1. Codes and their properties [18]
with descriptive statistic was conducted in 7 universities and 6 call centers in areas of Bangkok. About 800 questionnaire forms were distributed. Mainly, there were 2 parts to a questionnaire form that consisted of personal information and open questions for sentence listing. The questionnaire asked about the most frequently used sentences in daily general conversation, telephone conversation and news reports. Then, the answers would be processed, in order to obtain TTST, as in figure 3.

4. RESULTS

702 questionnaire forms had been returned but 34 returned forms were not answered in part 2 which was the part for sentence listing. The results from the survey can be presented in two parts as follows:

4.1. General Information from Participants

(a) Comparison between numbers of male VS female and comparison between numbers of general users VS heavy users

(b) Comparisons among ranges of age and comparisons among levels of education

![Figure 3. Overview of TTST Development](image)

![Figure 4. Pie charts for proportion of general information about participants (without consideration of missing values)](image)

4.2. Daily Used Thai Sentences (or Phrases)

3,467 sentences (or phrases) were obtained from 668 returned forms that had been completed, although some questions were not answered. After sorting by using possible and reasonable keywords to find 50 most frequently used sentences (or phrases) from the results, the list was generated. Some examples of the sentences (or phrase) are as follows:

![Table 2. First 7 most and last 6 of the most frequently used sentences (or phrases) in the list of TTST](image)

Note: the phonetic symbols in table 2 are the NECTEC used, which were defined in [23].

![Figure 5. Tonal analysis for TTST](image)

![Figure 6. Syllable analysis for TTST](image)

5. DISCUSSION

General information about the participants, according to figure 4 (a), found that there were more women than men, whereas numbers of general users were more than heavy users. From figure 4 (b), the majority of them are around 20-29 years old, whereas, almost 80% of them are representatives of good educated people in Bangkok and Thailand. The daily used Thai Sentences (or phrases) were generated from 50 most frequently used sentences (or phrase). As in table 2, each sentence (or phrase) meets the ITU-T Recommendation P.800 because each is simple, meaningful, short, and easy to understand. They were syllable analysis, to investigate if there are too short
or too long sentences (or phrases). Also, tonal analysis was conducted because tone is an important feature of Thai. Therefore, it could help to create each group of sentences (or phrases) to cover 5 tones when this TTST is applied to TSST. The data analyses are in the Figure 5-6. However, for some tones, numbers of sentences (or phrases) can vary because there is a different in tone between 准入 / ฉัน (kʰráp/kʰá) for male and female words, whereas, four third of number of sentences (or phrases) contains at least 3 tones and about three fifth of number of sentences (or phrase) consists of 4-6 syllables only.

6. CONCLUSION

This paper reveals some background information about Thai brains responding to Thai sounds that have a tonal feature, telecommunication technology, particularly VoIP. Also, looks at voice quality measurement and Thai resources, for both Thai text and Thai speech resources. However, this paper intends to present the approach using a survey, to avoid bias from researchers. This will lead to the creation of TTST which is the new Thai text resource for telephonometry before applying it to TSST and eventually supporting research in telecommunication areas in Thai environments. Hopefully, TTST, which is a part of TSST, will be in the NECTEC database as one of Thai resources for researchers. As for its contribution, the outcome from this will be useful for the advancement of telecommunication in Thailand, particularly VoIP quality measurement when VoIP services are used and voice quality is an important issue for users in the near future.

ACKNOWLEDGMENT

Thanks Mrs. Patcharika Choothrakul, Academic Staff, HLT, NECTEC, for Linguistics support and Mr. Gary Sherriff, International Coordinator, Faculty of Information Technology, KMUTNB, for editing, the reviewers for very good comments, and greatest gratitude to Dr. Gareth Clayton, co-author and beloved advisor of the 1st author, who has passed away.

REFERENCES