A Comparative Study of VoIP Quality Measurement from G.711 and G.729 Using PESQ and Thai Speech

Kiattisak Yochanang[1], Therdpong Daengsi[2], Tuul Triyason[3], and Pongpisit Wuttidittachotti[1]

[1] Faculty of Information Technology, KMUTNB, Bangkok, Thailand
{kiattisaky,pongpisitw}@kmutnb.ac.th
[2] Enterprise Services, JADS Comm Limited, Bangkok, Thailand
Therdpong1@yahoo.com
[3] School of Information Technology, KMUTT, Bangkok, Thailand
51501701@st.sit.kmutt.ac.th

Abstract. This paper presents the study of VoIP quality measurements from two popular codecs, G.711 and G.729, using the methods of Perceptual Evaluation of Speech Quality (PESQ) and Thai speech. In this study, from four lists of Thai speech, it has been found that G.711 provides better voice quality than G.729 in every condition of packet loss. Also, it has been found that Objective Listening Quality - Mean Opinion Score (MOS-LQO) of male speech is slightly higher than MOS-LQO of female speech, whereas MOS of child speech is the lowest. Then, MOS-LQO values from four Thai speech lists have been compared. Next, MOS-LQO from PESQ of male and female speech at the best condition have been compared with the Subjective Listening Quality Mean Opinion Score (MOS-LQS) from ACR listening tests in another laboratory. Lastly, referring to packet loss effects, objective MOS from PESQ have been compared with subjective MOS from conversation tests. It has been found that there is no significant difference among MOS-LQO from the four Thai speech lists, but it has been found that there is a significant difference between subjective MOS and objective MOS from each codec in each condition. Therefore, one can say that this is evidence that PESQ requires intensive study with Thai speech to modify PESQ for VoIP quality measurement in Thai environments confidently.

1 Introduction

VoIP quality measurement is one of the interesting issues referring to VoIP technology. Particularly, research on VoIP quality measurement based on Thai speech is very interesting because Thai is a tonal language, which is unique. At this moment, it still requires deep investigation, although some issues with the Thai language have been studied already, for example, subjective MOS studies with G-family codecs using subjective tests and Speex using PESQ [1-2].

Therefore, this study has been conducted in order to investigate VoIP quality measurement using PESQ with Thai speech, which has been applied to the Thai

B. Papasratorn et al. (Eds.): IAIT 2013, CCIS 409, pp. 242–255, 2013.

Speech Set for Telephonometry (TSST) [3], with unofficial collaboration between Faculty of Information Technology, KMUTNB, and School of Information Technology, KMUTT.

2 Background

VoIP quality measurement methods have been mainly classified into subjective methods and objective methods [4] [5]. Some of those are presented as follows:

2.1 Subjective Methods

Subjective result of voice quality is an important issue because it is necessary to be benchmarked for objective measurement tool calibration. It has been stated in [2] that, in general, it is agreed that subjective result of voice quality evaluation gives better reliability that the result from objective measurement.

Conceptually, subjective testing seems a very simple idea to obtain Mean Opinion Score (MOS), which is a key quality indicator for speech communication that is standardized by ITU-T [6]. MOS is the average of scores, using a 5-point scale (5=Excellent, 4=Good, 3=fair, 2=poor, 1=bad) from individual scores. However, for reliability of MOS, it requires 16 subjects or more, requiring controlled conditions (e.g., packet loss effects) in a quiet environment, such as a soundproof room [7].

Absolute Category Rating (ACR) testing is one of the better known subjective methods. This method has been issued by ITU-T; its result is a MOS of listening quality, which refers to how subjects rate what they hear during a call or speech sample [6-7]. It is used to obtain the absolute quality of the voice sample, through the direct hearing of the speech sample, without a reference speech sample [7]. Of course, ACR testing requires standard speech files; otherwise, variability in scores might be high. That means the MOS result might suffer from unreliability. Nevertheless, for its limitation, some effects such as delay effects cannot be tested using this method.

Therefore, conversation testing can be an alternative, particularly to evaluate conversational quality, which refers to how subjects rate the overall quality of a call impacted by, for example, echo and/or delay effects, thus it is the most realistic atmosphere recommended by ITU-T because this method can reach the same standard of realism [5] [7] [8]. Moreover, this method does not require speech files, while each round of testing can provide two scores. Therefore, with the same numbers of subjects, it should be able to decrease time of 50% to gather scores.

2.2 Objective Methods

Even though there is no objective measurement method that can replace subjective measurement method perfectly, objective measurement method has several advantages, such as, shorter time consumption, no collaboration with a lot of subjects and fewer endeavors [2, 9].

There are several objective measurement methods for speech communications, However, the major methods are PESQ for intrusive methods and E-model for non-intrusive methods, which have been confirmed by the search results from IEEE*Xplore*®, as shown in Table 1 [10] [11]. In that table, it can be seen from the search that the results of PESQ are higher than results of the E-model.

Table 1. Research directions using PESQ and E-model in last five years

Objective Method	Result	Remark
PESQ	41	Jan 2009 – Aug 2013
E-model	37	

For PESQ, which has been used in this study, it has been surveyed from [5] [12] [13] [14] [15] [16] and summarized in [17] as follows:

- PESQ is state-of-the-art in terms of objective voice quality measurement and has been claimed to have very high correlation with the subjective voice quality measurement methods
- it is the most common and popular method of intrusive measurement methods, including the original version, P.862 that supports narrow-band telephone networks and speech codecs, and P.862.2 which supports wideband telephone networks and speech codecs
- the original ITU-T P.862 PESQ supports only narrow-band telephone networks, whereas its new version was extended to support wideband telephone networks
- it uses the strength of both Perceptual Speech Quality Measurement (PSQM) and Perceptual Analysis Measurement System (PAMS), which are psycho-acoustic and cognitive models, and a time alignment algorithm respectively
- it is recommended to evaluate the impact of a codec to voice quality and to test the networks before operation
- it can be applied to evaluate several factors, such as transmission errors, codec errors, noise in the system, packet loss and time clipping
- it works as the model that compares the degraded signals to the original signals, instead of subjective evaluation
- this model permits the discovery of time jitter and identification of frames involved and which frames are affected by the delay and erased in order to prevent a bad score
- the average correlation between PESQ scores and the subjective scores was 0.935, reported in [14], whereas, it is claimed in [16] that the correlation is up to 0.95.

2.3 Thai Language and Thai Sound System

The Thai language is a tonal language used as the official language in Thailand that has population of over 65 million people. There are 44 Thai consonants that are classified into three groups known as High, Middle and Low class consonants.

The phoneme of Thai consists of 21 initial consonants, 11 cluster consonants, 9 final consonants, 21 vowels (consisting of short vowel, long vowel and diphthong) [18]. Particularly, there is a tonal feature that is very important for the Thai sound system, due to different tones of Thai speech which change the meaning of Thai words [18]. There are five tones shown in Table 2.

Moreover, from the research on neuroscience, it has been reported in [20] that from the study using Electroencephalogram technique and Low-Resolution Brain Electromagnetic Tomography (LORETA) with the Mismatch Negativity (MMN), with 9 Thai native listeners who do not have Chinese knowledge, that only the left hemisphere (LH) of Thai native speakers were activated significantly when listening to the condition with Thai speech which is their mother language, whereas, their right hemisphere (RH) were activated when listening to the condition with Chinese speech. Besides, it has been reported in [21] that the part of LH was predominant in the perception of the prosody of Thai speech sound, while the prosody of Chinese speech sound was dominated by part of the RH. That means only the left hemisphere can significantly respond to native speech sounds.

For speech samples that are necessary for the listening tests and different kinds of objective voice quality tests, ITU-T recommended that the speech samples must be meaningful and easy to understand [5], therefore, the 'Harvard sentences' that have been used widely cannot be used for voice or speech quality measurement in Thailand, due to those sentences are only in English [22].

Instead of using a set of Thai speech from the 'Multi-lingual speech database for telephonometry 1994' that's provided by NTT-AT with expensive costs [22], local Thai speech sets have been considered. However, instead of following [2] using the well-known Thai speech corpus 'LOTUS', which has been noted that many sentences are not simple when compared with the Thai Speech Set for Telephonometry (TSST) [3] [4]. TSST has therefore been considered for application, following [23] for further investigation. Some parts of speech sentences of TSST are shown in Table 3.

2.4 Related Works

In 2002, it was reported that results from the study with English and Dutch speech with loss effects using PESQ in [24] that the voice quality from male speech is higher than the voice quality from female speech.

Table 2. An example of Thai words with five tones

	Thai Tone				
	Mid	Low	Falling	High	Rising
Word	เลา	เหล่า	เล่า/เหล้า	เล้า	เหลา
Phonetic Symbol	lā	là	lâ	lá	lǎ
Meaning	noun classifier of a flute	set or group	to tell/alcohol	a (pig or duck) pen	to sharpen (a pencil)

Several years later in 2008, there was one research work from the same laboratory that is consistent with [24]. It presented voice quality with the GSM codec, English speech and six conditions [12]. It has been found that the voice quality from male speech using PESQ is higher than the voice quality from female speech. While it is inconsistent with [24], when using 3SQM, it was found that the voice quality from female speech is higher than the voice quality from male speech.

Similar to [25], the study with G.711 and G.729 using PESQ, it has been observed that the results from the experiment with delay effects of the voice quality from English female speech is slightly higher than the voice quality from English male speech. However, it has been observed from the same article that it is inconsistent with Chinese speech because the voice quality of Chinese male speech is slightly higher than the voice quality of Chinese female speech. Nevertheless, with loss effects, it has not been observed clearly from the experiment with G.711 about the differences of the voice quality from English male speech and English female speech, while the result of Chinese speech is consistent with the result of the experiment with delay effects.

For the study with Thai, both male and female speech, with Speex codec using PESQ [2], the result from the experiment with Thai speech is consistent with [24], and [12] that used PESQ tests. It has been observed that the voice quality from Thai male speech is higher than the voice quality from Thai female speech.

However, all above were mainly from objective tests using PESQ, the results are inconsistent with the result from the subjective tests, because it has been shown in [26] from interview tests, the voice quality from the female interviewer is slightly higher than the voice quality from the male interviewer. Besides, it has been compared between G.711 and G.729 referring to loss effects in [8] using conversation tests but it did not analyze type or gender of speech.

Therefore, it is still necessary to investigate and analyze Thai speech, both male and female.

3 Methodology

From Table 3, the speech lists have been adopted from [23], which follows [27] that recommended to have the length of each speech sample around 8-30 s, as shown in Table 4, as follows:

1) List2: there are three forms of speech samples (A, B and C). Each speech sample, with the length of 8 s, consists of two speech sentences. For each form, there are 2 child speech samples (a boy and a girl), 4 female speech samples and 4 male speech samples; therefore, totally there are 30 speech samples for this speech list.

2) List3: there are three forms of speech samples (A, B and C). Each speech sample, with the length of 8 s, consists of three speech sentences. Similar to List2, for each form, there are 2 child speech samples (a boy and a girl), 4 female speech samples and 4 male speech samples; therefore, totally there are 30 speech samples for this speech list.

3) List4: there are three forms of speech samples (A, B and C). Each speech sample, with the length of 10-12 s, consists of four speech sentences. Similar to List2, for each form, there are 2 child speech samples (a boy and a girl), 4 female speech samples and 4 male speech samples; therefore, totally there are 30 speech samples for this speech list.

Table 3. Selected Speech Samples for creating Subgroups of TSST as a Small Size Approach

Speech Group No.	Thai Sentences (or Phrases) with English Meaning
G1-1	สวัสดีครับ/ค่ะ (Hello.)
G2-1	หมายเลขที่ท่านเรียกไม่สามารถติดต่อได้ในขณะนี้ (The number you've tried to reach cannot be connected.)
G2-2	คิดถึงนะ (Miss you.)
G4-1	สอบถามข้อมูลอื่นเพิ่มเติมอีกไหมครับ/คะ (Would you like to ask for more information?)
G5-1	ต่อไปเป็นข่าวในพระราชสำนัก (Next, it is the royal news.)
G5-2	Where have you been? (ไปไหนมาเหรอ)
G6-1	กรุณาถือสายรอสักครู่ครับ/ค่ะ (Please hold on one moment.)
G6-2	ยินดีต้อนรับครับ/ค่ะ (Welcome.)
G7-1	วันนี้จะไปเที่ยวที่ไหนดี (Where are we going today?)
G7-2	ไม่ได้เจอกันตั้งนาน (Long time no see.)
G8-1	กำลังจะไปไหน (Where are you going?)
G8-2	ขอรบกวนเวลาสักครู่นะครับ/คะ (Please give me sometime.)
G9-1	วันนี้ไปกินข้าวที่ไหนกันดี (Where will we go to eat today?)
G9-2	จะกลับเมื่อไหร่ (When will you come back?)
G11-1	ดูแลรักษาสุขภาพด้วยนะ (Take care of your health.)
G11-2	ตกลงนะครับ/คะ (Are you ok?)
G12-1	กรุณาติดต่อกลับมาใหม่ (Please contact again.)
G12-2	กลับถึงบ้านหรือยัง (Have you reached home?)
G13-1	วันนี้เรียนวิชาอะไร (What subject are you going to study today? or What subject have you studied today?)
G13-2	จะกลับถึงบ้านกี่โมง (When will you reach home?)
G14-1	กำลังทำอะไรอยู่เหรอ (What are you doing now?)
G14-2	วันนี้รถติดมากเลย (Traffic is/was bad today.)
G15-1	ขณะนี้เวลาแปดนาฬิกา (Now, the time is eight a.m.)
G15-2	ขอโทษนะครับ/คะ (I'm sorry.)
G17-2	ยุ่งอยู่หรือเปล่า (Are you busy now?)
G18-2	ตื่นนอนหรือยัง (Have you woken up yet?)
G20-1	วันนี้อากาศร้อนมาก (Today is very hot.)
G22-2	กินข้าวหรือยัง (Have you eaten food?)
G23-1	ขอบคุณมากครับ/ค่ะ (Thank you very much.)
G24-1	นอนหลับหรือยัง (Have you slept yet?)
G25-2	แค่นี้ก่อนนะครับ/คะ (Ok, I have to hang-up now.)

Table 4. Lists of TSST that were created and selected for this study

Subset of TSST	Item	Length (s)	Speech group no.
List2	A	8	G12-2
			G20-1
	B	8	G6-1
			Subset of TSST
	C	8	G4-1
			G8-2
List3	A	8	G2-2
			G10-2
			G22-2
	B	8	G13-2
			G15-2
			G18-2
	C	8	G14-2
			G17-2
			G24-1
List4	A	10	G2-2
			G14-2
			G22-2
			G23-1
	B	10	G8-1
			G10-2
			G15-2
			G25-2
	C	12	G1-1
			G2-1
			G8-2
			G6-1
List10	A	30	G5-1
			G5-2
			G6-1
			G6-2
			G7-1
			G7-2
			G8-1
			G8-2
			G9-1
			G9-2
	B	30	G11-1
			G11-2
			G12-1
			G12-2
			G13-1
			G13-2
			G14-1
			G14-2
			G15-1
			G15-2

4) List10: there are two forms of speech samples (A and B). Each speech sample, with the length of 30 s, consists of ten speech sentences. Also, for each form, there are 2 child speech samples (a boy and a girl), 4 female speech samples and 4 male speech samples; therefore, totally there are 20 speech samples for this speech list.

Then, all speech lists, which consist of 110 speech samples, were tested using the PESQ tool and the network simulator in VoIP Lab, KMUTT with G.711 (A-law) and G.729, referring to packet loss effects (0%, 2%, 6% and 10%). Totally, there are 8 under test conditions (2 codecs x 4 loss conditions). Each speech sample was repeated 10 times per condition; therefore, the gathered data from this study are 8,800 records totally.

4 Results and Discussion

After obtaining the data and MOS-LQO [2], only comparison between MOS-LQO provided by G.711 and G.729 but also three issues have been considered, type of speech (child, female and male) and comparison between the results from G.711 and G.729, and the comparison among the results from four lists of speech samples. Moreover, the trends of results from PESQ tests were compared with the subjective results from the other study in another laboratory. The results are shown below:

4.1 Comparison of MOS from Thai Speech of Child, Female and Male

For overall, as shown in Fig. 1 – Fig. 4, referring to packet loss effects, MOS- LQO from G.711 is higher than G.729. It is consistent with the subjective result as reported in [8]. Then the type of speech has been compared, as shown in Fig. 1 – Fig. 4, it can be seen that:

1) At the best condition of packet loss (0%), the MOS-LQO from G.711 no significant difference among child speech, female speech and male speech was found. Whereas, it has been discovered that there is a slight difference among three of them from G.729, for male speech over female speech over child speech.

2) For overall, it can be seen that MOS-LQO from G.711 is better than MOS-LQO from G.729 in each condition under testing.

3) At the packet loss of 2%, 6% and 10%, the MOS-LQO from both G.711 and G.729, differences among three of them from G.729, for male speech over female speech over child speech were discovered.

4) The results from all speech lists tend to be in the same manner that MOS-LQO from male speech is higher than female speech, whereas MOS-LQO from child speech is the worst. This might be one of the reasons that ITU-T does not provide child speech samples in [28]. For overall, the result from this study is consistent with the results from PESQ tests in [24] [12], whereas, it is inconsistent with the observed results from 3SQM tests in [12] and PESQ tests in [25].

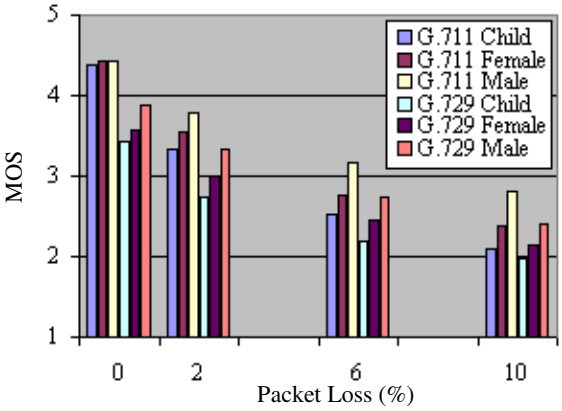

Fig. 1. Comparison of MOS among child speech, female speech and male speech from List2

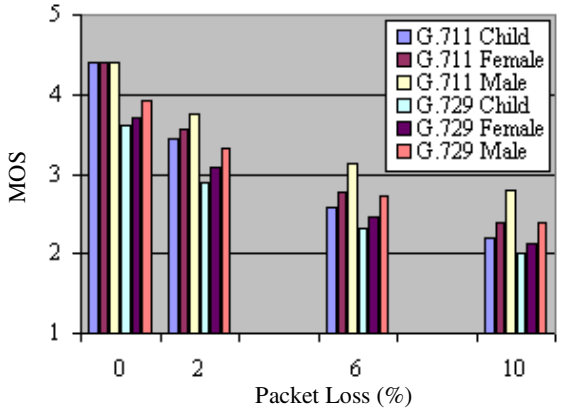

Fig. 2. Comparison of MOS among child speech, female speech and male speech from List3

Moreover, it is inconsistent with the result from [26]. It has been reported that interviewees prefer female interviewer over male interviewer, which might stem from the gender of speech.

Therefore, from the results, it can be summarized that PESQ tool responds to low speech frequency over high speech frequency because fundamental frequency of male speech is lower than female speech and child speech, as presented in [9]. This is evidence which can be used to improve objective tools when working with male and female speech. Of course, at present, child speech samples should not be used when studying objective methods (e.g. PESQ).

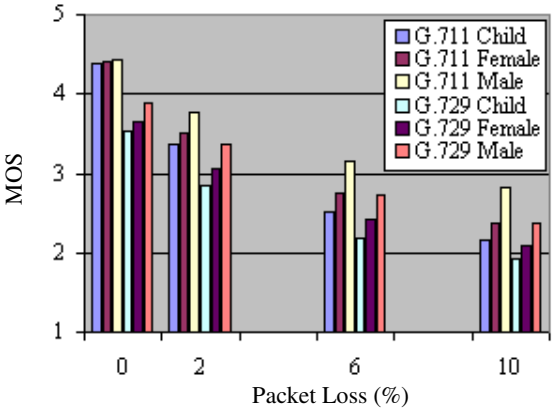

Fig. 3. Comparison of MOS among child speech, female speech and male speech from List4

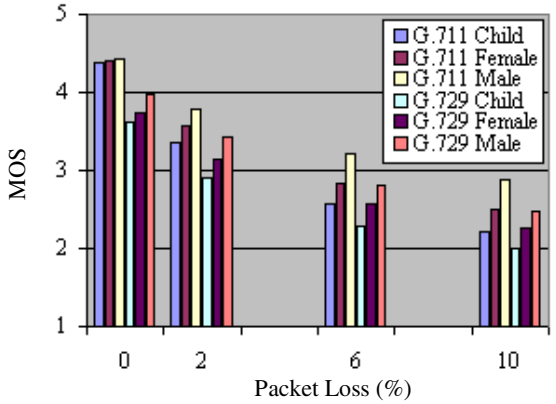

Fig. 4. Comparison of MOS among child speech, female speech and male speech from List10

4.2 Comparison of MOS from Four Speech Lists

As shown in Table 5, the results without MOS-LQO from child speech, one can see no difference of MOS-LOQ provided by G.711 from four speech lists, the maximum difference is 0.02 at packet loss of 2%, whereas MOS-LQO is exactly the same, 2.59, at packet loss rate of 10%. While, the standard deviation tends to be higher when the packet loss rate is increased.

For the results from G.729, as shown in Table 6, although there is no significant difference among the four speech lists, it can be observed that MOS-LQO values from List10 in all conditions are higher than List2 –List4, whereas, the standard deviation values from List10 are lower than the others.

Therefore, it can be summarized that each Thai speech list from TSST can be used with PESQ tool with confidence.

Table 5. Comparison of MOS-LQS values among four speech lists with G.711

Loss (%)	MOS				Standard Deviation			
	List2	List3	List4	List10	List2	List3	List4	List10
0	4.42	4.41	4.42	4.42	0.06	0.06	0.05	0.05
2	3.66	3.66	3.64	3.64	0.30	0.30	0.27	0.27
6	2.96	2.95	2.95	2.95	0.33	0.33	0.35	0.35
10	2.59	2.59	2.59	2.59	0.31	0.32	0.34	0.34

Table 6. Comparison of MOS-LQS values among four speech lists with G.729

Loss (%)	MOS				Standard Deviation			
	List2	List3	List4	List10	List2	List3	List4	List10
0	3.76	3.82	3.77	3.85	0.19	0.15	0.15	0.14
2	3.17	3.21	3.21	3.29	0.31	0.26	0.28	0.19
6	2.56	2.60	2.57	2.70	0.28	0.26	0.26	0.18
10	2.22	2.25	2.24	2.37	0.25	0.25	0.25	0.17

4.3 Comparison of Objective MOS and Subjective MOS

Similar to Section 4.2, MOS-LQO values from child speech were discarded, MOS-LQO from PESQ of male and female speech at the best condition have been compared with MOS-LQS from ACR listening tests in another laboratory [9], it has been found that MOS-LQO of 4.41-4.42, as in Table 5, is slightly higher than MOS-LQS of 4.23 for G.711, whereas MOS-LQS of 4.18 is higher than MOS-LQO of 3.76-3.85, as in Table 6, for G.729.

Without available subjective MOS from listening tests referring to packet loss effects, the result of objective MOS, MOS-LQO, from G.711 and G.729 in this study with Thai speech has been compared with the result of subjective MOS, MOS-CQS [1], from conversation tests that are available in [8].

As shown in Fig. 5, G.711, MOS-CQS is lower than MOS-LQO at the best condition of packet loss. However, MOS-CQS is higher than MOS-LQO at the other conditions, particularly packet loss of 6% and 10%.

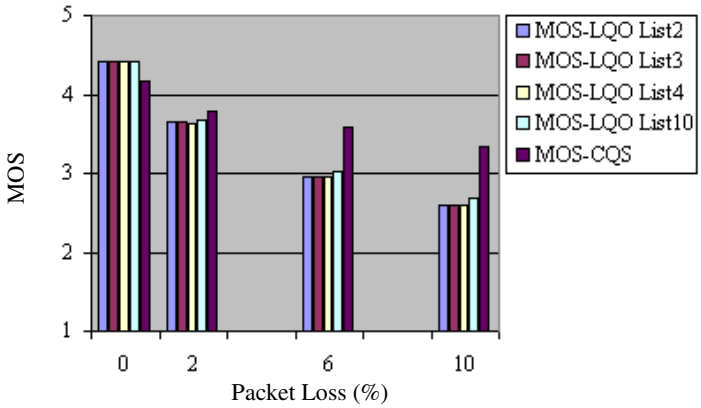

Fig. 5. Comparison of MOS-LQO and MOS-CQS from G.711

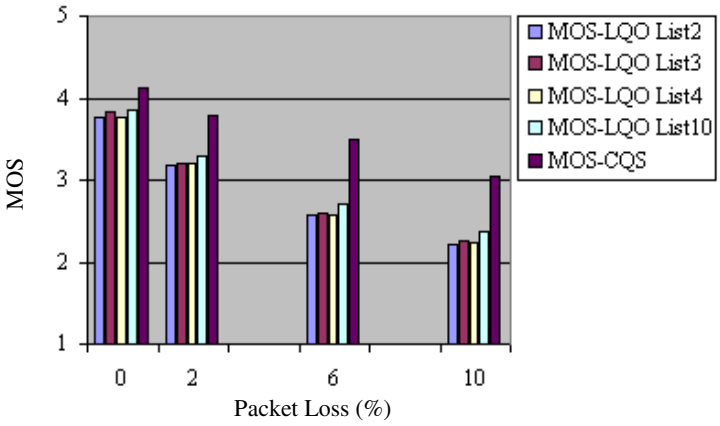

Fig. 6. Comparison of MOS-LQO and MOS-CQS from G.729

For G.729, as shown in Fig. 6, it can be seen that MOS-CQS at each condition it is higher than MOS-LQO significantly, particularly at packet loss of 2%, 6% and 10%.

However, this issue should be re-investigated and compared to MOS-LQS and MOS-LQO using Thai speech. Then, the localization-mapping factor based on Thai users should be found for PESQ calibration.

5 Conclusion and Future Work

This paper presents the comparative study of VoIP quality measurement from G.711(A-law) at 64 Kbps and G.729 at 8 Kbps using PESQ and Thai speech, referring to packet loss effects.

From this study with Thai speech, firstly, it has been found referring to packet loss that MOS-LQO from G.711 is slightly higher than MOS-LQO from G.729 in each condition. Also, it has been found that MOS-LQO from Thai male speech is higher than MOS-LQO from Thai female speech. Whereas, MOS-LQO from Thai child speech is the lowest, compared to the others. Therefore, it can be summarized that PESQ tends to give higher values of MOS-LQO for speech sample that have low fundamental frequency like male speech, due to the fundamental frequency of male speech is lower than female and child speech. Thus, PESQ should be improved to include this issue as future work.

Secondly, it has been found that there is no significant difference from this study that MOS-LQO from four speech lists (List2, List3, List4 and List10), although MOS-LQO from List10 seems slightly higher that MOS-LQO from others. Therefore, each of them can be used with confidence.

Lastly, it has been found that there are significant differences between objective MOS from PESQ and subjective MOS from conversation tests. Therefore, the MOS-LQS should be conducted using the four speech lists in this paper and then compared with MOS-LQO from this study as future work, in order to modify PESQ for Thai speech especially.

Acknowledgments. Thank you to the VoIP Lab, School of Information Technology, KMUTT for the PESQ tool and the network simulator used to gain these results. Thank you so much Mr. Gary Sherriff, the international coordinator, Faculty of Information Technology for editing.

References

1. Daengsi, T., Wutiwiwatchai, C., Preechayasomboon, A., Sukparungsee, S.: A Study of VoIP Quality Evaluation: User Perception of Voice Quality from G.729, G.711 and G.722. In: Proc. 9th IEEE-CCNC 2012, Las Vegas, NV, pp. 342–345 (2012)
2. Triyason, T., Kanthamanon, P.: Perceptual Evaluation of Speech Quality Measurement on Speex Codec VoIP with Tonal Language Thai. In: Proc. 5th International Conference IAIT, Bangkok, Thailand, pp. 181–190 (2012)
3. Daengsi, T., et al.: The Development of a Thai Speech Set for Telephonometry, http://desceco.org/O-COCOSDA2010/proceedings/paper_53.pdf (accessed November 2012)
4. Daengsi, T., Preechayasomboon, A., Sukparungsee, S., Wutiwiwatchai, C.: Thai Text Resource: A Recommended Thai Text Set for Voice Quality Measurements and Its Comparative Study. KKU Science Journal 40(4), 1114–1127 (2012)
5. ITU-T Recommendation P.800 (Methods for subjective determination of transmission quality), http://www.itu.int/rec/T-REC-P.800-199608-I/en (accessed November 2012)
6. Telchemy. Voice Quality Measurement (Application Note), http://www.telchemy.com/appnotes/TelchemyVoiceQualityMeasurement.pdf (accessed August 2013)
7. Rango, F.D., et al.: Overview on VoIP: Subjective and Objective Measurement Methods. IJCSNC 6, 140–153 (2006)
8. Daengsi, T., Sukparungsee, S., Wutiwiwatchai, C., Preechayasomboon, A.: Comparison of Perceptual Voice Quality of VoIP Provided by G.711 and G.729 Using Conversation-Opinion Tests. IJCIM 20(1), 21–26 (2012)
9. Daengsi, T., Wutiwiwatchai, C., Preechayasomboon, A., Sukparungsee, S.: VoIP Quality Measurement: Insignificant Voice Quality of G.711 and G.729 Codecs in Listening-Opinion Tests by Thai Users. Information Technology Journal 8(1), 77–82 (2012)
10. http://ieeexplore.ieee.org/search/searchresult.jsp?queryText%3DVoIP+Quality+E-model&addRange=2009_2013_Publication_Year&pageNumber=1&resultAction=REFINE (accessed August 2013)
11. http://ieeexplore.ieee.org/search/searchresult.jsp?queryText%3DVoIP+Quality+PESQ&addRange=2009_2013_Publication_Year&pageNumber=1&resultAction=REFINE (accessed August 2013)
12. Goudarzi, M., Sun, L.: Performance analysis and comparison of PESQ and 3SQM in live 3G mobile networks, http://www.tech.plym.ac.uk/spmc/staff/mgoudarzi/Performance%20analysis%20and%20comparison%20of%20PESQ%20and%203SQM.pdf (accessed August 2013)
13. ITU-T Recommendation P.862. Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone (2001)
14. ITU-T Recommendation P.862.2. Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs (November 2007)

15. Karapantazis, S., Pavlidou, F.-N.: Voip: A comprehensive survey on a promising technology. Computer Networks 53, 2050–2090 (2009)
16. Ditech Networls. Limitations of PESQ for Measuring Voice Quality in Mobile and VoIP Networks, https://www.google.co.th/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDEQFjAA&url=http%3A%2F%2Fwww.packagingdigest.com%2Ffile%2F4534-white_paper_on_the_limitations_of_PESQ.pdf%3Fforce%3Dtrue&ei=NTdFUoXmHIHQrQfU2oHoDw&usg=AFQjCNGKfvZ5aWrk_kD2wr8IkamNfku5zw&bvm=bv.53217764,d.bmk (accessed September 2013)
17. Daengsi, T.: VoIP Quality Measurement: Recommendation of MOS and Enhance Objective Measurement Method for Standard Spoken Thai Language. Ph.D. Thesis, Faculty of Information Technology, King Mongkut's University of Technology North Bangkok (November 2012)
18. Sodanil, M., Nitsuwat, S., Haruechaiyasak, C.: Thai Word Recognition Using Hybrid MLP-HMM. IJCSNS 10(3), 103–110 (2010)
19. Wutiwiwatchai, C., Furui, S.: Thai speech processing technology: A review. Speech Communication 49, 8–27 (2007)
20. Sittiprapaporn, W., Chindaduangratn, C., Kotchabhakdi, N.: Brain electric activity during the preattentive perception of speech sounds in tonal languages. SJST 26, 439–445 (2004)
21. Sittiprapaporn, W., Chindaduangratn, C., Kotchabhakdi, N.: Functional Specialization of the Human Auditory Cortex in Processing of Speech Prosody: A Low Resolution Electromagnetic Tomography (LORETA) Study, http://anchan.lib.ku.ac.th/kukr/bitstream/003/17786/1/KC4205011.pdf (accessed December 2012)
22. IEEE. Recommended practice for speech quality measurements. IEEE Trans. on Audio Electroacoust. 17(3), 225–246 (1969)
23. Suayroop, K., et al.: A VoIP Quality Measurement Study with Thai Speech Sets Using PESQ and G.711A-law. In: Proc. 34th Electricall Engineering Conference (EECON-34), Chonburi, Thailand (2011)
24. Sun, L., Ifeachor, E.: Perceived Speech Quality Prediction for Voice over IP-based Networks. In: Proc. IEEE Int. Conf. Communications 2002 (ICC 2002), New York, NY, pp. 2573–2577 (2002)
25. Ren, J., Zhang, H., Zhu, Y., Goa, C.: Assessment of effects of different language in VOIP. In: Proc. Int. Conf. Audio, Language and Image Processing (ICALIP 2008), Shanghai, China, pp. 1624–1628 (2008)
26. Daengsi, T., Wutiwiwatchai, C., Preechayasomboon, A., Sukparungsee, S.: Speech Quality Assessment of VoIP: G.711 VS G.722 Based on Interview Tests with Thai Users. IJITCS 4(2), 19–25 (2012)
27. ITU-T Recommendation P.862.3 (Application guide for objective quality measurement base on Recommendations P.862, P.862.1 and P.862.2) (November 2007)
28. Annex B Speech files, ITU-T Recommendation P.501, http://www.itu.int/net/itu-/sigdb/genaudio/AudioForm-.aspx?val=10000501 (accessed August 2013)