A Personalized Learning Platform to Improve English Pronunciation at Word Level for Thai EFL Learners based on End-to-End Automatic Speech Recognition

Kongpop Boonma, Phongsatorn Ousakulwattana,
Science Classroom in University Affiliated School (SCiUS),
PSU Wittayanusorn Surat Thani School,
Thailand,
Nattapol Kritsuthikul,
Language and Semantic Technology Laboratory (LST),
National Electronic and Computer Technology Center (NECTEC), Thailand,
Jirapond Muangprathub,
Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Thailand
Tirapond Jaroensak,
International College,
Prince of Songkla University, Surat Thani Campus, Thailand

***Abstract*— ASR (Automatic Speech Recognition) is favorably chosen as a learning technology, which is used for English pronunciation practice. This research aims to build a personalized learning platform to improve English pronunciation at the word level for Thai EFL (English as a Foreign Language) learners by using ASR to detect mispronounced sounds. ASR models are built with an End-to-End learning approach with a Thai-English mispronounced words dataset. The practice of English pronunciation particularly focuses on eleven problematic consonant sounds of Thai EFL students according to the previous studies of English pronunciation in Thai contexts. These eleven consonant sounds are divided into five groups: 1) /ð/-/θ/-/tθ/, 2) /ʒ/-/ʃ/, 3) /dʒ/-/tʃ/, 4) /z/-/s/ and 5) /b/-/p/. The five of Grade 12 Thai Students who are native Thai speaker were selected as sampling process. The pre and post-tests results show that the samples have the most problem with /ð/-/θ/-/tθ/ (29%) followed by /b/-/p/ (22%), /dʒ/-/tʃ/ (22%), /z/-/s/ (18%) and /ʒ/-/ʃ/ (9%) respectively. In conclusion, this study reveals that 60% of the samples have improved their pronunciation after using our system.**

***Index Terms*— English as a Foreign Language (EFL), Mispronunciation, Thai, Automatic Speech Recognition (ASR), End-to-End, Personalized Learning Platform, Computer aided pronunciation training (CAPT)**

I. Introduction

Today, English plays a big role as an international language and becomes a dominant language for intercultural communication. One of the most important language skills is speaking skills. Similarly, English is taught as a foreign language in Thailand, but Thai EFL students still have a very low levels of English proficiency [1].

The type of English that we speak doesn't matter very much as far as we speak in an intelligible way. If you live in a country where there is no traditional use of English and no people who speak it for general communication purposes, the English pronunciation you are going to speak may reflect the distinction between your native language and English.

Furthermore, the English pronunciation that can be understood in your home country may not be the case in another. Though English is not the media for communication in Thailand, sometimes Thai people use borrowed English words, but pronounced in Thai ways that native speakers may not understand [2].

There are vary of English pronunciation practice software. Most of software can detect the right sound but a few can detect the mispronunciations sound. Hence, this paper aims to design a system as a self-learning material for Thai EFL students to practice their English pronunciation by build an

Kongpop Boonma and Phongsatorn Ousakulwattana are with Science Classroom in University-affilaited School program, PSU Wittayanusorn Surat Thani, Surat Thani, Thailand, e-mail: {kongpopboonma1912, phongsatorn2004}@gmail.com

Nattapol Kritsutikul is with Language and Semantic Technology Laboratory (LST), National Electronics and Computer Technology Center (NECTEC), Thailand email: nattapol.kritsuthikul@gmail.com

Tiraporn Jaroensak is with International College, Prince of Songkla University, Surat Thani Campus, Thailand email: tiraporn.j@psu.ac.th

Jirapond Muangprathub is with Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Thailand email: jirapond.m@psu.ac.th

algorithm with the ASR to detect mispronunciation sounds. The ASR is defined as a cutting-edge technology that allows a computer or even a hand-held PDA to identify words that are read aloud or spoken into any sound-recording device. The ultimate purpose of ASR technology is to allow 100% accuracy with all words that are intelligibly spoken by any person regardless of vocabulary size, background noise, or speaker variables. For this research, we use the ASR as a tool to detect mispronunciation sounds. The content in Introduction will be discussed more in Literature review.

II. Literature review

A. *English Pronunciation Problems in Thais*

English is a pluricentric language; that is to say, varieties of English are emerging with different marked accents. These accents become remarkable in global communication when English plays an important role as a lingua franca. English learners from different parts of the world produce different sounds influenced by their mother tongue, and that possibly causes them a barrier and misunderstanding. Therefore, an international intelligibility relies on the clarity of English pronunciation and utterances, rather than native-like pronunciation [3]. A number of previous studies report problems of English pronunciation found in Thai EFL learners that Thai English learners still have encountered difficulties in mispronunciation of some sounds such as /s/, /z/, /θ/, /r/, /l/, /v/ and /f/ [4] [5] [6]. In addition, Jaroensak & Saraceni reported that some distinctive features of Thais' English pronunciation found in lingua franca communication were a shift of initial and final consonant sounds. The consonant sound /t/ was shifted into /tʃ/ and the sound /tʃ/ was pronounced as /ʃ/ [7]. These mispronunciations possibly cause a barrier to interlocutors' intelligibility; on the other hand, it leads to speakers' loss of confidence and anxiety to speak English.

B. *Overview of Using a Software as an English Language Learning Material on Thai EFL Learners*

Studies on improving English pronunciation skills by using the software have already been widely conducted. The study of Sudrutai [8] enhanced the English pronunciation competence of Thai students by using Speech Analyzer software. This tool gives the visualizations of the raw waveform and intensity of the sound wave of the voice record of both the student and the native speaker, hence they can compare on the contrast in their pronunciations between themselves and the native speaker. As an example, that some of the words or sound is considered to force the sound through the oral cavity differently from the language so the students could know how to control their oral muscles and air pressure from the vocal tract more certainly. The results reported that using the speech analysis software could significantly enhance the students' pronunciation.

Jeereapan [9] created Detect Me English application, The English correction software to analyze the phonological sounds produced by the students, to create awareness of incorrect pronunciations in Thai Elementary students. The results showed that the sample groups were not aware of many of the English phonological rules.

Parthanasin & Blackford [10] created a case study of Siri in iPad as voice recognition for English Pronunciation practice. The results of the analysis clearly prove that Siri application could recognize utterances spoken by a native speaker of English better than those spoken by a non-native speaker in all categories. It can be claimed that correct pronunciation is essential for the machine to recognize an utterance.

Steven Graham [11] used SpeaKIT's speech recognition software to study whether this software can actually improve English primary school English language learners' skills. The results showed that it is apparent that audio-visual speech recognition has the potential to have something for every child and that it is up to the teacher to identify the student's individual learning style and adapt the implementation of the program according to the needs of the students in their specific classroom.

C. *End-to-end Speech Recognition*

The speech research community found reliable and effective approach [12] both in lab prototype [13] and industry prototype [14]. The major advantages of End-to-End approach are reduced intensive workload of the audio engineer robustness acoustic models. By comparison with the traditional speech system based on Hidden Markov Models (HMMs), End-to-End approach required a few inputs e.g. (1) raw audio and (2) transcription of the raw audio input, and so on. One disadvantage of End-to-End approach is that they require high performance computing power especially GPUs at each step.

End-to-End ASR rely on sophisticated pipelines composed of multiple algorithms and fine-tuning processes by the audio engineer. To achieve the ASR tasks, an End-to-End ASR pipeline was constructed from various concepts and tutorials e.g., Very Deep Learning and M5 Network [15], Deep Speech [16], Speech Command Classification with TorchAudio [17], and Thai Speech Command Recognition with TorchAudio [18]. The pipeline allows us to apply the ASR as the mispronunciation detection tool. In overall, the pipeline applied very deep convolutional neural networks (CNNs), up to 34 weight layers, to processing the raw audio data as inputs. The M5 network architecture used to filter the raw data in into the receptive fields around 20 ms each. This size is similar to speech processing applications that often use receptive fields ranging from 20 ms to 40 ms for training and testing the network. Finally, we construct the model by applying the pipeline in the training process based on the Thai mispronunciation corpus.

III. System Architecture

In this section, we will describe the proposed system and system overview for our proposed system, an E-learning platform.

A. *System Requirements and Design*

The proposed system aims at improving English pronunciation for Thais by detecting the mispronounced sounds from the learner and giving the correct pronunciation and advice to them. By doing that, the learner would know what sounds they speak wrong so they could correct those wrong sounds to improve their English pronunciation. Fig. 1 shows the system architecture that outlined each major components for the proposed system.

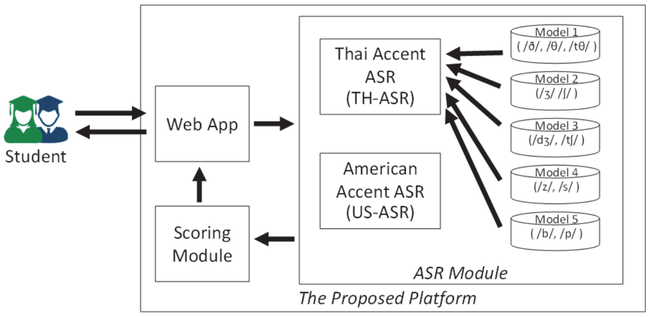


Fig. 1. System Architecture

The system is developed as a web-based application using HTML5, CSS3, ECMAScript 2020 as front-end framework, Flask, a Python micro web framework, as back-end framework in advantage of convenient of Python libraries invocation, and MySQL as a relational database management system (DBMS) to store All the system information, e.g. learning activities, score, and so on. The system consists of two modules which are ASR Module and Scoring Module.

ASR Module is used to predict voice audio from the student to the system by either uploading voice audio files or directly streaming audio into the module. There are two types of ASR, American Accent ASR powered by Google speech-to-text API, and ASR for five mispronunciation model as mentioned in the Model Training Process.

Scoring Module is used to manage the student learning path based on the student performance interaction in the system. Fig. 2 shows the scoring process, started by the user have to speak the word from the system, the system then detects the user's speech with Google speech-to-text API to match the user's speech with the word user speak. If the Google speech-to-text API output matches with the word user speak, we consider that user has no problem with that sound, but if Google speech-to-text API output is not equal to the spoken word.

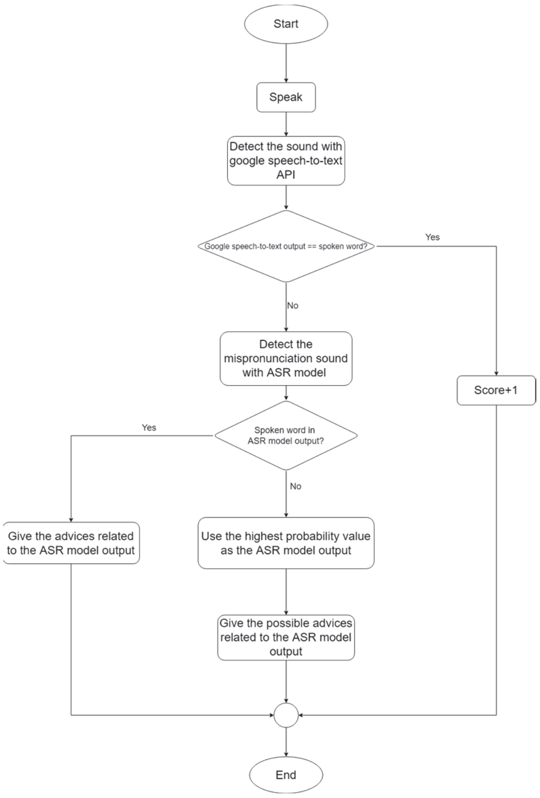

Start
Speak
Detect the sound with google speech-to-text API
Google speech-to-text output == spoken word?
Yes
No
Detect the mispronunciation sound with ASR model
Score+1
Spoken word in ASR model output?
Yes
No
Give the advices related to the ASR model output
Use the highest probability value as the ASR model output
Give the possible advices related to the ASR model output
End

Fig. 2. Scoring process

The system is going to detect mispronounced sounds with 5 ASR models. If ASR predicted word output equal to spoken word, the output is sent to the user to suggest the correct pronounced sounds. But if ASR predicted word output is not equal to spoken word, the output is going to be the highest probability value from 5 ASR models. For the process above, the process can allocate the data adequately for indicating the student's mispronouncing. The system will generate suggestions to the student as shown in Fig. 3, 4, and 5.

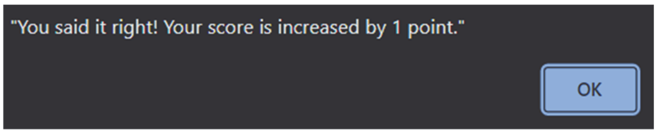

"You said it right! Your score is increased by 1 point."

Fig. 3. An example of no mispronunciation

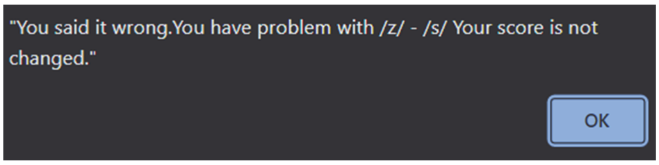

"You said it wrong.You have problem with /z/ - /s/ Your score is not changed."

Fig. 4. An example of /z/ - /s/ mispronunciation found

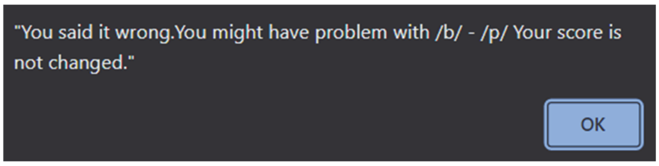

"You said it wrong.You might have problem with /b/ - /p/ Your score is not changed."

Fig. 5. An example of /b/ - /p/ mispronunciation found

Finally, both ASR Module and Scoring Module are store the processing results into the database. The database consists of 4 tables which are `user`, `word_list`, `score`, and `asr` as shown in Fig. 6. The `user` table stores user data. The `word_list` table stores English words used in the pronunciation practice. The `score` table stores the pronunciation score which is divided into 5 groups according to the English dataset group. The `asr` table stores ASR process data.

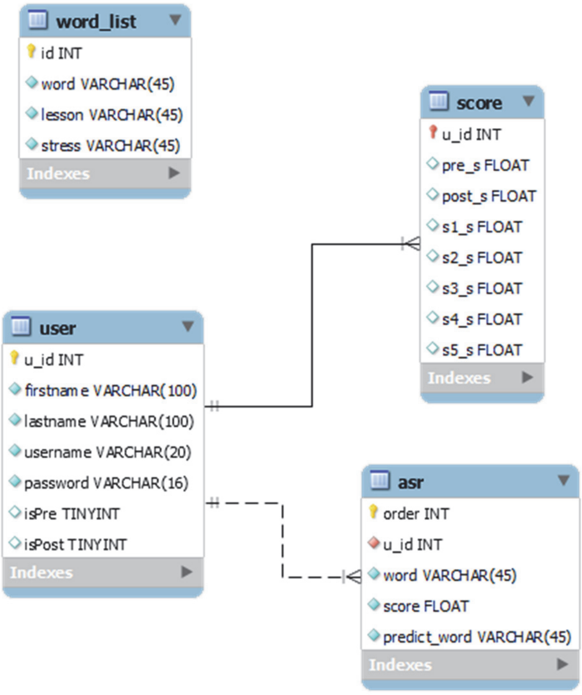


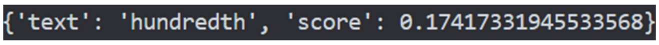
Fig. 6. Database Diagram

B. Audio Acquisition

From our study, the sounds that have problems in English pronunciation for Thai people have come up with 11 sounds, namely /ð/, /θ/, /tθ/, /ʒ/ /ʃ/, /dʒ/, /tʃ/, /z/, /s/, /b/ and /p/. We record those sounds by native Thai students, undergraduates, and teaching professionals in English and non-English subjects. The record is created for Thai-English mispronounced pronunciation corpus to train the ASR model to detect mispronounced sounds. The corpus is divided into 5 dataset 1) /ð/, /θ/, /tθ/, 2) /ʒ/ /ʃ/, 3) /dʒ/, /tʃ/, 4) /z/, /s/ and 5) /b/, /p/.

C. Model Training Process

To build the model, TorchAudio [19], a library for audio and signal processing with PyTorch, is introduced for building ASR models to detect mispronounced sounds. The audio data have to be set in 2-channel audio. The training data file consists of 2 classes 1) audio filename and 2) transcript for each audio in Comma-separated values (CSV) file format. The model gives an output consisting of 2 data. 1) a predicted word and 2) a confident value. The predicted word is the word that the model predicts from the voice audio. The confidence value is the probability of the predicted word. Please note that this kind of value from Google Speech to Text ASR is not available due to the limitation of free version of the API. The output will be used as a source to indicate whether the student pronounces this word correctly or not. An example is shown in Fig. 7.



```
{'text': 'hundredth', 'score': 0.17417331945533568}
```

Fig. 7. Output of ASR model

At the end of this process, the five models are conducted according to the five datasets that described in the Audio Acquisition section.

IV. Experiment

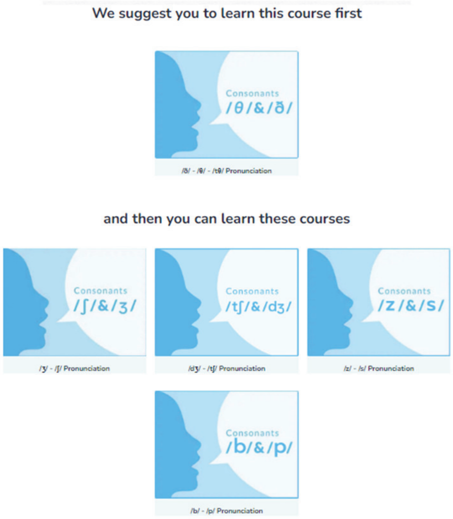
We test the system with 5 native Thai speaker students studied in PSU Wittayanusorn Surat Thani School. Testers have to test with our learning process start by pre-test. After finishing the pre-test, testers have to learn the lessons order by lesson recommendation. Finally, Testers have to do the post-test to see the differences after using the system. Post-test words are all similar to pre-test words.

The learning starts on the pre-test. The pre-test has 15 words divided into 5 groups according to datasets. There are 3 words for each group as shown in Table I.

TABLE I
Group of words in for pronunciation practice

Group	Word
/ð/-/θ/-/tθ/	altogether
	smooth
	though
/ʒ/-/ʃ/	bashful
	chicanery
	distinguish
/dʒ/-/tʃ/	charge
	fragile
	indulge
/z/-/s/	decease
	overseas
	practice
/b/-/p/	belonging
	bilingual
	dabble

After finishing the pre-test, the system prepares the lesson according to the user's score. The lesson is also divided into five groups based on the five datasets as shown in Fig 8.



We suggest you to learn this course first

and then you can learn these courses

Fig. 8. Lesson recommendation related to pre-test score

V. Results

The bar graph in Fig. 9 shows that 60% of testers improved after using our system. Nevertheless, the highest score learner could do is only 5 out of 15 points.

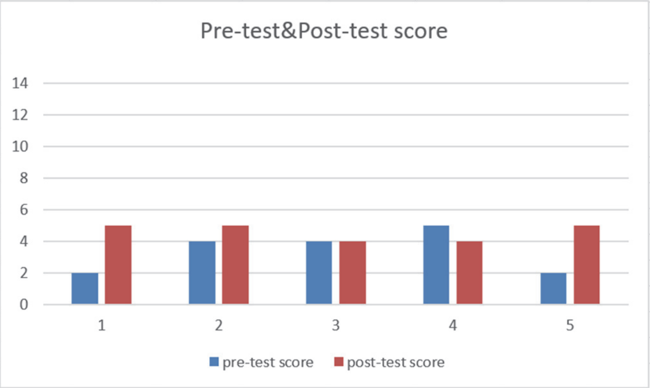


Fig. 9. Pre-test & Post-test score

Table II shows the Word error rate (WER) of each ASR model. We can see that all models have low accuracy. The reason why all models have low accuracy is that our corpus has not enough amount of data to train the appropriate accuracy ASR model.

TABLE II
WER of each ASR

Model	WER
/ð/-/θ/-/tθ/	44%
/ʒ/-/ʃ/	60%
/dʒ/-/tʃ/	56%
/z/-/s/	29%
/b/-/p/	33%

We also found that there are 6 words that none of the samples could speak correctly which are altogether, smooth, though, fragile, decease, belonging, and dabble as shown in Table III and IV.

TABLE III
PRE-TEST PRONUNCIATION RESULT

Pre-test						
word	Learner 1	Learner 2	Learner 3	Learner 4	Learner 5	count
altogether	0	0	0	0	0	0
smooth	0	0	0	0	0	0
though	0	0	0	0	0	0
bashful	1	1	1	1	0	4
chicanery	1	0	0	1	0	2
distinguish	0	1	1	1	1	4
charge	1	0	0	0	1	2
fragile	0	0	0	0	0	0
indulge	0	1	0	0	1	2
decease	0	0	0	0	0	0
overseas	1	0	0	0	0	1
practice	1	1	1	1	1	5
belonging	0	0	0	0	0	0
bilingual	1	0	1	1	1	4
dabble	0	0	0	0	0	0

TABLE IV
POST-TEST PRONUNCIATION RESULT

Post-test						
word	Learner 1	Learner 2	Learner 3	Learner 4	Learner 5	count
altogether	0	0	0	0	0	0
smooth	0	0	0	0	0	0
though	0	0	0	0	0	0
bashful	1	1	1	1	0	4
chicanery	0	1	1	1	0	3
distinguish	1	1	1	1	1	5
charge	0	1	0	0	1	2
fragile	1	0	0	0	0	1
indulge	1	0	0	0	1	2
decease	0	0	0	0	0	0
overseas	1	0	0	0	0	1
practice	1	0	1	1	1	4
belonging	0	0	0	0	0	0
bilingual	1	1	0	0	1	3
dabble	0	0	0	0	0	0

The experimental result shows that samples have problem with /ð/-/θ/-/tθ/ the most (29%) followed by /b/-/p/ (22%), /dʒ/-/tʃ/ (22%), /z/-/s/ (18%), /ʒ/-/ʃ/ (9%) according to Google speech-to-text outputs as shown Fig. 10.

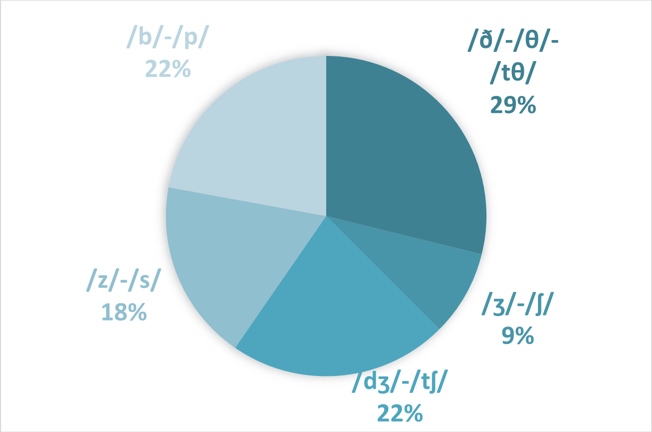


Fig. 10. Percentage of mispronunciation result detecting by Google speech-to-text API

Fig. 11 shows the percentage of mispronunciation output from 5 ASR models. The highest mispronunciation output count is ʒ/-/ʃ/ (61%) followed by /b/-/p/ (11%), /z/-/s/ (11%), /ð/-/θ/-/tθ/ (10%) and /dʒ/-/tʃ/ (7%). However, this data is still unreliable because low accuracy models provide many wrong outputs.

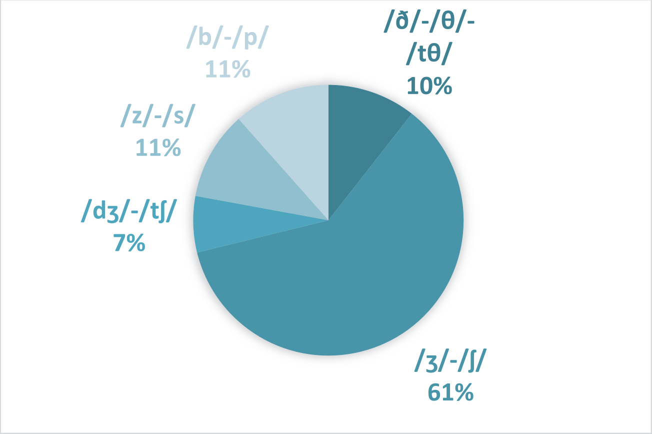


Fig. 11. Percentage of mispronunciation output from 5 ASR models

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a prototype of a personalized learning platform to improve English pronunciation at the word level by using ASR to detect mispronounced sounds and give correct pronunciation. Our study aims at 11 consonant sounds that Thai EFL learners have problems with (see Table I)

Based on the experimental results, it can be concluded that using ASR to detect mispronounced sounds has a potential to improve pronunciation skills. Moreover, we also found that samples have problems with /ð/-/θ/-/tθ/ the most (29%) followed by /b/-/p/ (22%), /dʒ/-/tʃ/ (22%), /z/-/s/ (18%), /ʒ/-/ʃ/ (9%).

Since the accuracy of the prototype depends on the ASRs, we intend to employ better ASR models. We also intend to find more detection method of mispronunciation for Thais in future work. Furthermore, samples and the amount of words on the experiment are still too low to provide reliable results. Therefore, in further research, we want to apply more samples and words to extend experimental data for reliable results. The corpus is also needs to expand to provide proper accuracy. We will build the model by applying multi-class classification training approach so that we can merge the five ASR models into a single model.

ACKNOWLEDGMENT

This project was supported by Science Classroom in University Affiliated School (SCiUS) with Prince of Songkla University Surat Thani Campus. The funding of SCiUS is provided by the Ministry of Higher Education, Science, Research and Innovation, Thailand. We would like to thank Phongsatorn Ousakulwattana and the other 16 people for helping on collect the data for the Thai-English accent corpus and Tankwan Khongcharoen for helping.

REFERENCES

[1] EF EPI 2021, "EF English Proficiency Index", https://www.ef.com/wwen/epi/, 2021.

[2] Wei, Y., "Insights into English Pronunciation Problems of Thai Students", 2002.

[3] Jenkins, J., "A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language", *Applied linguistics*, *23*(1), 2002, 83-103.

[4] Moxon, S., “Exploring the Effects of Automated Pronunciation Evaluation on L2 Students in Thailand”, *IAFOR Journal of Education*, *9*(3), 2021, 41-56.

[5] Sahatsathatsana, S., “Pronunciation problems of Thai students learning English phonetics: A case study at Kalasin University”, *Journal of Education, Mahasarakham University*, *11*(4), 2017.

[6] Wei, Y., & Zhou, Y., “Insights into English Pronunciation Problems of Thai Students”, 2002.

[7] Jaroensak, T., & Saraceni, M., “ELF inThailand: Variants and Coinage in Spoken ELF in Tourism Encounters”, *REFLections*, *26*(1), 2019, 115-133.

[8] Arunsirot, S., “Implementing a Speech Analyzer Software to Enhance English Pronunciation Competence of Thai Students”, 2017.

[9] Phomprasert, J., “Creating Awareness of Incorrect English Pronunciation in Thai Elementary School by using the Detect Me English Application”, 2020.

[10] Pathanasin, S., & Blackford, M., “Voice Recognition for English Pronunciation Practice: A Case Study of Siri in iPad”, 2015.

[11] Graham, S., “An Exploratory Case Study to Investigate Perceived Pronunciation Errors in Thai Primary School Students Using Audio-Visual Speech Recognition, Teaching English with Technology”, 2020.

[12] Jinyu Li., "Recent Advances in End-to-End Automatic Speech Recognition," *arXiv preprint arXiv: 2111.01690v2*, 2022.

[13] S. Watanabe, T. Hori, S. Kim, J. R. Hershey, and T. Hayashi, "Hybrid CTC/Attention Architecture for End-to-End Speech Recognition," *IEEE Journal of Selected Topics in Signal Processing*, 11(8), 2017, 1240–53.

[14] J. Li, R. Zhao, Z. Meng, Y. Liu, W. Wei, S. Parthasarathy, V. Mazalov, Z. Wang, L. He, S. Zhao, et al., "Developing RNN-T Models Surpassing High-Performance Hybrid Models with Customization Capability," *in Proc. Interspeech*, 2020, 3590–4.

[15] Wei Dai, Chia Dai, Shuhui Qu, Juncheng Li, and Samarjit Das., "Very Deep Convolutional Neural Networks for Raw Waveforms," *arXiv preprint arXiv:1610.00087*, 2016.

[16] Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, et al., "Deep speech: Scaling up end-to-end speech recognition," *arXiv preprint arXiv:1412.5567*, 2014.

[17] Alexis Gatignol., “Speech Command Classification with TorchAudio,” `https://github.com/pytorch/tutorials/blob/master/intermediate_source/speech_command_classification_with_torchaudio_tutorial.py`, 2022.

[18] Workshop on NLP/AI R&D in iSAI-NLP-AIoT 2021., “Thai Speech Command Recognition with TorchAudio,” `https://colab.research.google.com/drive/1ensKfWzt6WEvmAZTrtMtyUrX1i5JBkMk#scrollTo=yUHwDMnYqI8l`, 2021.

[19] Yao-Yuan Yang, Moto Hira, Zhaoheng Ni, Anjali Chourdia, Artyom Astafurov, Caroline Chen, Ching-Feng Yeh, Christian Puhrsch, David Pollack, Dmitriy Genzel, Donny Greenberg1, Edward Z. Yang, Jason Liany, Jay Mahadeokar, Jeff Hwang, Ji Chen, Peter Goldsborough, Prabhat Roy, Sean Narenthiran, Shinji Watanabe, Soumith Chintala, Vincent Quenneville-Bélairy, and Yangyang Shi., "TorchAudio: Building Blocks for Audio and Speech Processing," *arXiv preprint arXiv:2110.15018v2*, 2022.

Kongpop Boonma is studying Grade 12 in Science Classroom in University-affiliated school program (SCiUS), PSU Wittayanusorn Surat Thani School. He has been internship student in Language and Semantic Techology Laboratory, NECTEC, Thailand.

Phongsatorn Ousakulwattana is studying Grade 12 in Science Classroom in University-affiliated school program (SCiUS), PSU Wittayanusorn Surat Thani School. He has been internship student in BIOTEC, Thailand.

Nattapol Kritsuthikul received the B.S. degree in Computer Science from Bangkok University in 2000. He received the M.S. degrees in Information Technology from the King Mongkut's Institute of Technology Ladkrabang in 2006. He received the Ph.D. degree in Information Science from the Japan Advanced Institute of Science and Technology (JAIST) in 2017. He is currently the Artificial Intelligence in Education researcher of the Language and Semantic Technology (LST) Research Team at NECTEC in Thailand.

Tiraporn Jaroensak is a lecturer in Global Englishes and Research Methodology in Social Sciences at International College, Surat Thani Campus, Prince of Songkla University. She has her Ph.D. from the School of Languages and Applied Linguistics, University of Portsmouth, UK. Her doctoral research focused on English as a lingua franca (ELF) in tourism contexts. Her study explored pragmatic strategies in meaning negotiation. Accordingly, her research interests are ELF, the pragmatics of ELF, and implications for English Language Teaching.

Jirapond Muangprathub (nee Tadrat) received the B.Sc. degree in Applied Mathematics from the Prince of Songkla University, Thailand, in 2002, and the M.Sc. and Ph.D. degrees in Computer Science from King Mongkut's Institute of Technology (KMITL), Bangkok, Thailand, in 2005 and 2011, respectively. In 2011, she joined the Department of Applied Mathematics and Informatic, Faculty of Science and Industrial Technology, Prince of Songkla University (PSU), Surat Thani Campus, Surat Thani, as a Lecturer. She has been with the Department of Applied Mathematics and Informatic, PSU, where she was an Assistant Professor in July 2014. Her current research interests include data analysis, knowledge-bases system, knowledge representation, data mining, information retrieval, artificial intelligent, formal concept analysis, rough set theory, case-based reasoning, image processing, and internet of things.