Proposed of Using Speech Recognition Technology to Detect Read Aloud in Thai Tone Indications for Primary Education Students

Wudhijaya Philuek[a]*

[a]* *Assistant Professor, Ph.D., Department of Innovation and Educational Technology, Faculty of Education, Nakhon Sawan Rajabhat University, E-mail: wudhjijaya@nsru.ac.th*

Article Info
Received 23 July 2023
Revised 8 September 2023
Accepted 11 December 2023
Available online December 2023

Abstract

This academic article has objectives were 1. to study the techniques in data analyzing of speech recognition systems techniques to detect Thai Tone Indications for primary education students, and 2. to study the result of speech recognition systems techniques to detect in Thai Tone Indications for primary education students. Sample group in this research were 30 primary education students in Grades 4 – 6 of Wat Na Klang School in Nakhon Sawan Province, Thailand. Research tools was speech recognition systems techniques to detect reading aloud skills in Thai tonal conjugation.

Results showed that,

1.the three techniques involved in sound analysis consisted of Spectrogram Sound Wave Change, Wavelet Transform Audio Format Conversion, and two sets of data similarity comparison techniques named Dynamic Time Warping (DTW). The researchers used three techniques in analysis and comparison for making the more similar to the sound signal than using a single technique in analytics.

2. the result of using speech recognition system to detect reading aloud in Thai Tone Indications showed results of pronunciation practice and compared the similarity of audio signals that after using the technique in 5 analytical times, the largest average result was 93.52 Percent, the least average result was 87.75 Percent, and the total average result was 90.66 Percent.

Keywords: Sound Wave, Thai Tone Indications, Speech Recognition, Primary Education Students

Introduction

At present, Information Technology is playing a wide role in various fields, especially education by applying in various ways including the use of computer-assisted teaching, mixed media systems, information systems, database systems, and the Internet, etc. Due to the impact of advances in Information Technology, the format or method of educational management has changed from the traditional education that was based on teachers or teachers as the center of learning. knowledge to individual education, focusing on technology to promote self-learning, emphasizing morality and ethics (Malaiwong, 1997). Speech Recognition is a technology that allows a computer to function like a human cognition process and to correctly identify interests.

For example, it enables a computer to recognize information from a sound and a computer digitizes it which can be analyzed or processed (Korpunyapipat, 2015)

Learning any language, whether a native language or a foreign language, starts with learning the listening and pronunciation skills first, and whether language acquisition is successful or not depends on your ability to lead. to use effectively able to communicate successfully in the use of language for communication, pronunciation is very important (Naiyaphat, 2005). Thai language is the official language that all elementary school students must read and write according to the indicators and learning content in the core learning of the Basic Core Curriculum, B.E. 2551, presented that "Thai language is a national identity, a cultural treasure that creates unity, and enhances the personality of the nation to be Thai person, and also a tool for seeking information and experiences to develop knowledge as well as to use in career development is a valuable treasure worthy of preservation and continuation of the Thai nationality forever" (Ministry of Education, 2020). Thai language has contents, namely linguistic rules that language users must know and use the language correctly. In addition, literature and literature, as well as children's play songs Nursery rhymes, riddles, folk songs, folk literature Be part of the culture Which is valuable to the study of Thai language, so must study literature. language wisdom that convey feelings, thoughts, values, traditions The story of the past society and the beauty of language. in prose poems Poetry to appreciate the pride of what the ancestors have accumulated and inherited until now (Phamontri, 2015). However, it must be admitted that Thai educational policies do not have preparations or plans to deal with various emergency situations when comparing normal classroom teaching with the form established by the state. that it is not as effective as learning in the classroom because each learner is given the opportunity to access different educational information, which is in line with Panto (2020) saying that poor learners The family did not have enough income to support the educational materials. Or some schools are located in remote rural areas without access to the Internet or electricity, thus causing these children to lose learning opportunities. In addition, the Secretariat of the Education Council (2023) also stated that most of the learners lack parents. In providing advice or monitoring, because parents have to go out to work outside, while their children have to study by themselves at home, may make students not interested in their studies in the form held. Boonnithi (2021) said that the academic performance or chievement of students often depends on reading ability, especially reading spelling of students in the first grade, which is an important basis for starting reading practice to developing advanced reading skills all from reading spelling. almost all words The teachers are very necessary to develop activities and teaching styles to promote the spelling of students with problems with problems that have higher problems because when students can read, they will send to other skills as well. To help develop reading and spelling abilities, there are many ways that teachers need to find a process that best suits the context and problems of the students.

Thai letters, although they are fixed sound marks but there are many words that have problems with how to pronounce them. In the practice of reading Thai words, if the learner does not have the correct skills and understanding of tonal conjugations, it will cause misunderstandings and confusions, resulting in miscommunication or misunderstandings due to changes in the language. Tonal sounds in reading have the effect of changing the meaning of words (Na Nakorn, 1981). Hence, learning Thai language of Thais' Primary Education Students showed the significance that most of them have problems in Reading and Writing especially Tonal Diversion and Tonal Writing (Watcharasukhum, 2013; Thodthankun and Chansuwan, 2016; Dipatee and Narapongkasem, 2018; Cholkerkkait, 2018; Poosawad, 2018; Mardyusoh and Tornram, 2020)

The importance above shows that in learning management, teachers or instructors need to train learners to develop more clear practical skills through practicing pronunciation skills and learning skills of students. Students must be studied as a guideline for future student development. The researcher therefore would like to investigate the model and techniques in speech recognition to enhance practice of reading aloud in Thai Tone Indications for young learners

Objective

1. to study the techniques in data analyzing of speech recognition systems techniques to detect Thai Tone Indications for primary education students.
2. to study the result of speech recognition systems techniques to detect in Thai Tone Indications for primary education students.

Literature Review

In order for us to correctly conjugate Tone Marks or Intonation Marks, we must first understand the different parts of Tone Marks or Intonation Marks. In word formation, the key components of a word are initial consonants and vowels which are the most elementary component in some words, there may be spelling and orthography adds the beginning. However, the indispensable is the consonants, there are 44 Thai consonants according to the main textbooks in Thai which dividing these consonants into 3 groups, called “Three in Arithmetic” or “Tayyang” (Thonglor, 1972) that explaining as follows:

Three in arithmetic is a group of letters that are separated into groups of 44 consonants, namely High Consonants, Middle Consonants, and Low Consonants.

There are 11 High Consonants letters: ข ฃ ฉ ฐ ถ ผ ฝ ศ ษ ส ห

There are 9 Middle Consonants letters: ก จ ค ต ฎ ฏ บ ป อ

There are 24 Low Consonants letters: ค ฅ ฆ ง ช ซ ฌ ญ ฑ ฒ ณ ท ธ น พ ฟ ภ ม ย ร ล ว ฬ ฮ

The classification of consonants into 3 groups is based on the sound is important that is any consonant has not been conjugated with a high accent is classified as a high-voiced letter. Any consonant that has not been conjugated into a low accent is classified as a low-voiced consonant. We use all 3 groups of tone indications or tonal conjugations, dividing the tones into 2 types:

1. Tonality has a form which refers to the tonal tone that has a clear pitch mark above the letter. There are 4 forms: low tone, falling tone, high tone and rising tone. Tonal tone is called low tone, falling tone, high tone and rising tone, respectively such “ครุ่น” “คลื่น” “เกล้า” “ใกล้” “จ๊ะ” etc.

2. Formless tones is sounds that have a high and low melody according to the group of Thai characters formless tonal form can be read aloud, such as “นา” “หนะ” “นาก” “นะ” “หนา” etc.

From the above, there are 5 tonal in Thai language: mid tone, low tone, falling tone, high tone and rising tone, respectively, we use tones to distinguish the meaning of words in Thai as can be seen from the following example which not cover in all form of tones.

Example

[See “สี”]+ mid tone = ‘Color สี’

[See “สี”]+ low tone = ‘Four สี่’

[See “สี”]+ falling tone = ‘tooth ซี่’

[See “สี”]+ high tone = ‘To die (in Chinese meaning) ซี้’

[See “สี”]+ rising tone = ‘Color สี’

Zsiga and Nitisaroj (2007) studied the acoustic phonetics of Thai tonal sounds and concluded that Ordinary tone (M) is characterized by the middle pitch. The low tone (L) is characterized by descending from the lowest beginning to the lowest end of the pitch. The falling tone (F) is characterized by a high pitch, then clearly moving up and down. The high tone (H) has a high pitch beginning and then rises with a slight curve. The rising tone (R) tone begins at a low pitch, then curves down to near the bottom of the pitch and then curves upward.

The research on the acoustic phonetics of tone indications in Thai is very useful in further developing the speech recognition system and pronunciation training for foreigners including pronunciation training for people with pronunciation impairment, such as research by Nasani (2003) and Suriyasophapan (2003) showed that People without larynx should practice pronouncing Master and Junior tones, etc.

In speech processing, speech recognition is applications to analyze utter of speech which commonly used by researchers (Kershet and Bengio, 2009) Speech recognition is concerned with the words being spoken (Kikel, 2019) and is useful for people with various group of people, such as those with problems in speaking who find reading the words difficult or who have difficulties recognizing and spelling words, such as dyslexia people (Hanifa, Isa and Mohamad, 2017). Since speech recognition deals with converting audio into text, its effectiveness depends heavily on the language and the text corpus (Sharma, 2019). On the other hand, speaker recognition is to identify the person who is speaking. Pitch, speaking style, and accent are some of the features that contribute to the differences (Kikel, 2019). Speaker recognition technology has been used in various applications, such as biometric, security, and even human-computer interaction. Speech recognition in terms of recognition, purpose (in recognizes what is being said and converts them into text), focus (to identify and digitally record what the speaker is saying and Vocabulary of what is being said by the speaker and turns the words into digital texts), and application (Speech to text) (Biometric Today, 2018).

There are many algorithms for speech recognition, Dynamic Time Warping or DTW is one of algorithm based on dynamic programming which provides a temple matching for different lengths of pronunciation. It is a nonlinear warping technique and used more widely that requires less data and can be used as the most proficient and simple speech recognition algorithm (Zhi-Qiang et. al. 2019).

Research Methodology

This research titled the effects of effects of using speech recognition technology techniques to detect read aloud in Thai tone indications or tonal conjugations of young learners. There are operating procedures as follows:

1. Procedures for conducting research

1.1 Survey to collect information on problems in learning Thai tone indications or tonal conjugations. by using a questionnaire on the opinions of teachers and students

1.2 Learn about speech recognition technology

1.3 The sound obtained from the data collection of the sample was analyzed and compared with the conversion of each sound technique.

1.4 Apply each technique to analyze the results.

2. Scope of research

2.1 Demographic scope and target group in collect sound data.

2.1.1 The population used in the research was primary school students of Wat Na Klang School Academic Year 2021

2.1.2 The target group used in this research were 30 students in grades 4 - 6 at Wat Na Klang School (in Nakhon Sawan Province, Northern Part of Thailand), academic year 2021.

2.2 Scope of content

The data used to survey the students' problems with reading aloud in Thai tone indications or tonal conjugations consisted of words that were tonal formless, the tonal sounds match the tonal form, and the tone that does not match the tonal form in total of 100 words, and researchers collected sound data using 35 words that the students often mispronounce which categorized into 15 formless tones, 12 tones matching the tonal form, and 12 tones. whose sound does not match the tonal form 8 words.

3. Tools used in conducting research

3.1 Working data

3.1.1 The student's sound data obtained by collecting sound data using 35 words that students often mispronounced.

3.1.2 Data Analysis Program

3.2 Working equipment

3.2.1 Computer

3.2.2 Mobile phone

3.2.3 Techniques used in data analysis

3.2.3.1 Spectrogram technique

3.2.3.2 Wavelet Transform Technique

3.2.3.3 Dynamic Time Warping (DTW) Technique

3.2.4 Microsoft Excel

3.2.5 Any Video Converter Ultimate Program

3.2.6 Free MP3 Cutter Program

3.2.7 Praat Program

3.2.8 MATLAB program

3.2.9 Audio data

4. Method of collecting data

Step 1: Explore the problem of tonal words that are often mispronounced.

Step 2: After surveying, collect sound data using 35 words that students often mispronounce.

Step 3: To collect audio data, researchers asked for cooperation from students in the Thai language department. There was 1 male and 1 female each to collect the prototype sound data and collect the sample group's sound data by practicing the pronunciation for a total of 5 times.

Step 4: Analyze the sound data and compare it with the original sound to find a prototype in the same age range of students in Grades 4 - 6 before analyzing the data.

5. Data analysis

Researchers conducted the research in work steps are as follows:

Step 1: Understanding the Problem, researcher has studied the relevant research and has studied the problem of reading aloud and Thai tone indications or tonal conjugations, and study techniques used for analyzing various audio signals, Sound Wave Dynamics, Spectrograms, Wavelet Transforms, Analogy of Sequences that differ in time or speed, Dynamic Time Warping (DTW).

Step 2: Data Understanding, researchers studied the available data and considers the feasibility of data analysis. In collecting the data, researchers surveyed the students' problems with reading aloud and changing the Thai tone indications or tonal conjugations as in Table 1, Table 2 and Table 3 as follows:

Table 1: A Tonal Formless

Thai Words			**Tone indications**
คา (Ka)	มา (Ma)	เดือน (Duan)	Mid Tone (Tone 1)
ขาด (Kard)	จาก (Jak)	หลบ (Lop)	Falling Tone (Tone 2)
ฆาต (Kard)	แรด (Rad)	พูด (Pood)	Falling Tone (Tone 3)
รัก (Rak)	พัก (Pak)	ฟัด (Fad)	High Tone (Tone 4)
หนา (Nar)	ขา (Kha)	สวย (Suay)	Rising Tone (Tone 5)

Table 2: A tone that matches the tonal form

Thai Words			**Tone indications or (Writing Form)**	**Tone indications (Sound Form)**
ข่า (Kha)	สง่า (Sa-Nga)	หมี่ (Mee)	Falling Tone (Tone 2)	Falling Tone (Tone 2)
ก้อง (Kong)	จ้า (Jar)	ป้า (Par)	Falling Tone (Tone 3)	Falling Tone (Tone 3)
โต๊ะ (Toh)	แก๊ส (Gas)	กั๊ก (Kak)	High Tone (Tone 4)	High Tone (Tone 4)
แต๋ว (Taew)	จ๋า (Jar)	ปิ๋ม (Pim)	Rising Tone (Tone 5)	Rising Tone (Tone 5)

Table 3: A tone that do not matches the tonal form

Thai Words				**Tone indications (Writing Form)**	**Tone indications (Sound Form)**
ว่า (War)	ค่า (Kar)	ง่า (Ngar)	โล่ง (Long)	Falling Tone (Tone 3)	Falling Tone (Tone 3)
ย้อน (Yorn)	พ้อง (Pong)	ฟ้า (Fah)	ค้ำ (Kam)	High Tone (Tone 4)	High Tone (Tone 4)

Step 3: Data preparation is the importing of raw data that has been converted into data that will be used in the next step, as shown in Figure 1.

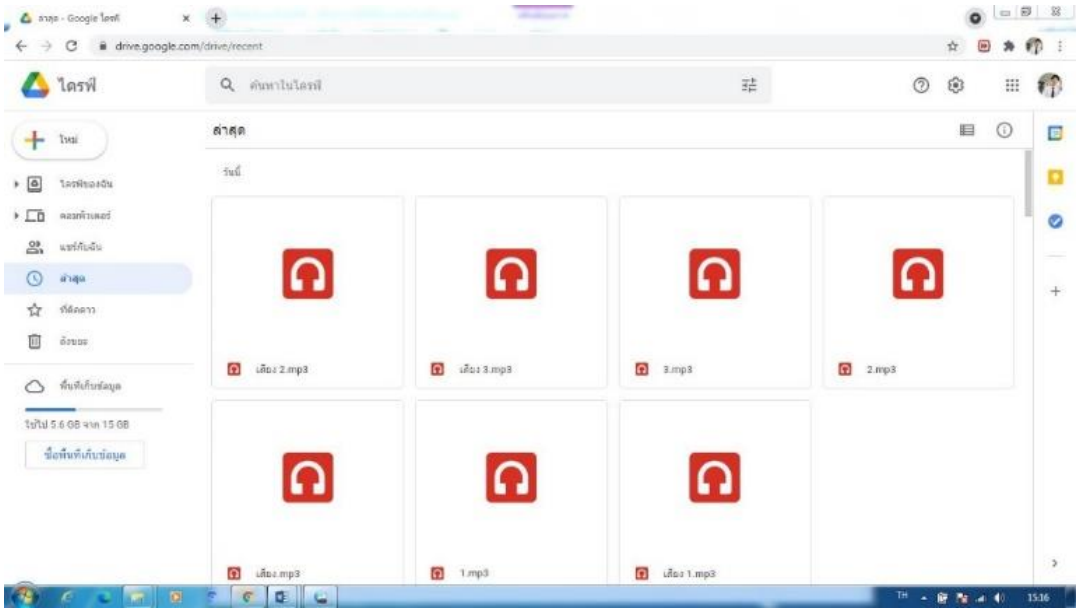


Figure 1: Sample of audio data from data collection

Selecting the appropriate variables for analysis and converting all variables to the same format to prepare the data for modeling can be divided into 3 sub-steps as follows:

1. Convert the resulting audio data into a .wav audio file using Any Video Converter Ultimate as shown in Figure 2.

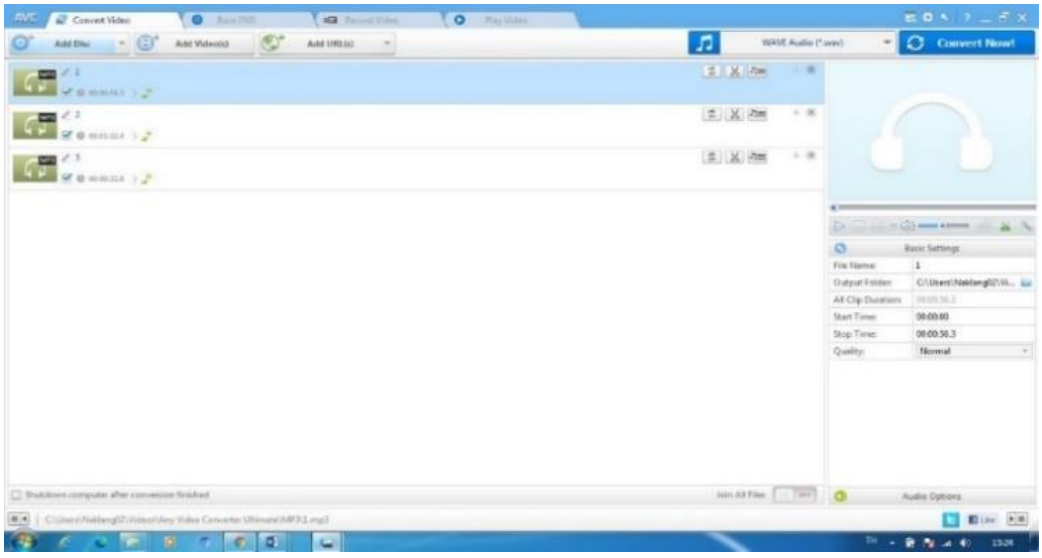


Figure 2: Convert the audio to .wav file format.

2. Cut the converted audio into words by the researcher using Free MP3 Cutter program as shown in Figure 3.

UBRU International Journal Ubon Ratchathani Rajabhat University
Vol.3 No.3 September - December 2023

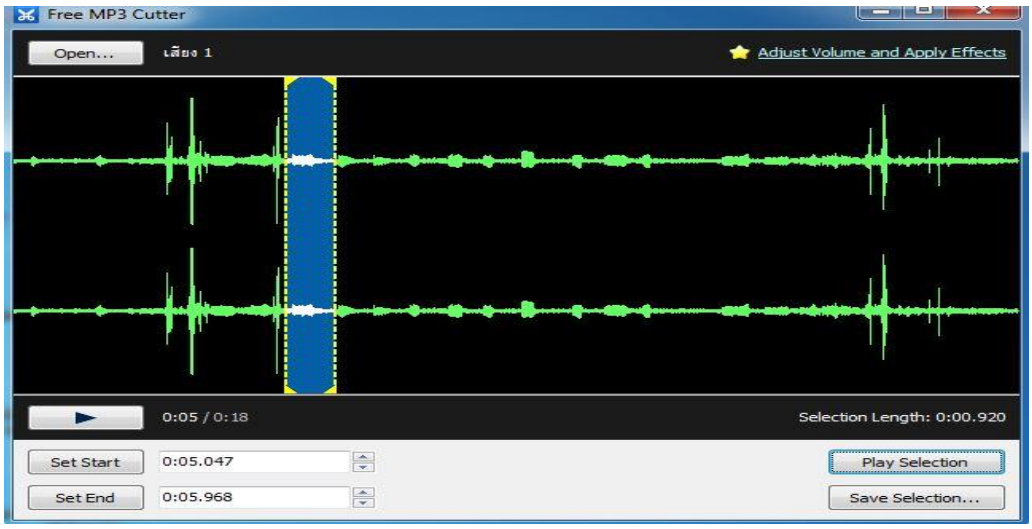


Figure 3: Cutting audio into words

3. Grouping the tonal words in each form and the sounds that have been cut into words and analyzed as shown in Figure 5.

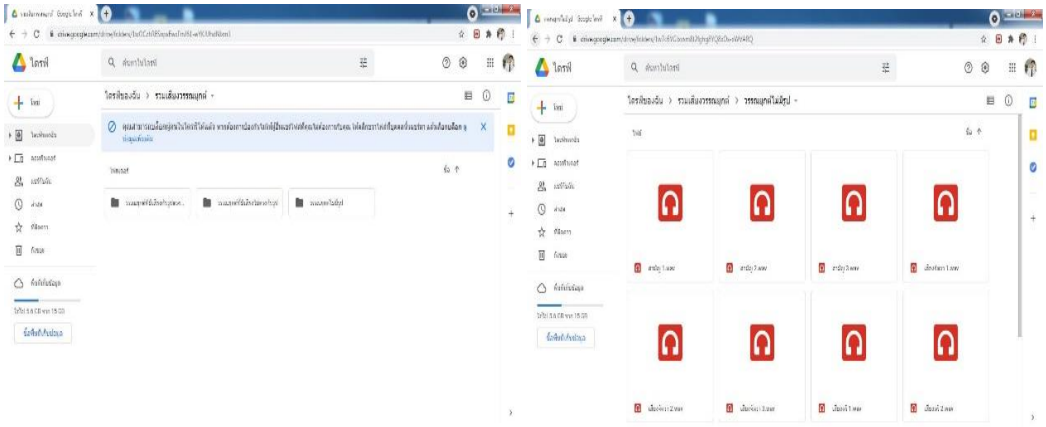


Figure 4: The tonal groups in each picture and the sounds that have been cut into words.

Step 4: Data analysis is the use of transcoded and cut-to-word audio data to compare sound similarity analysis with the following steps.

1. The researcher chose Praat and MATLAB programs for data analysis as shown in Figure 5.

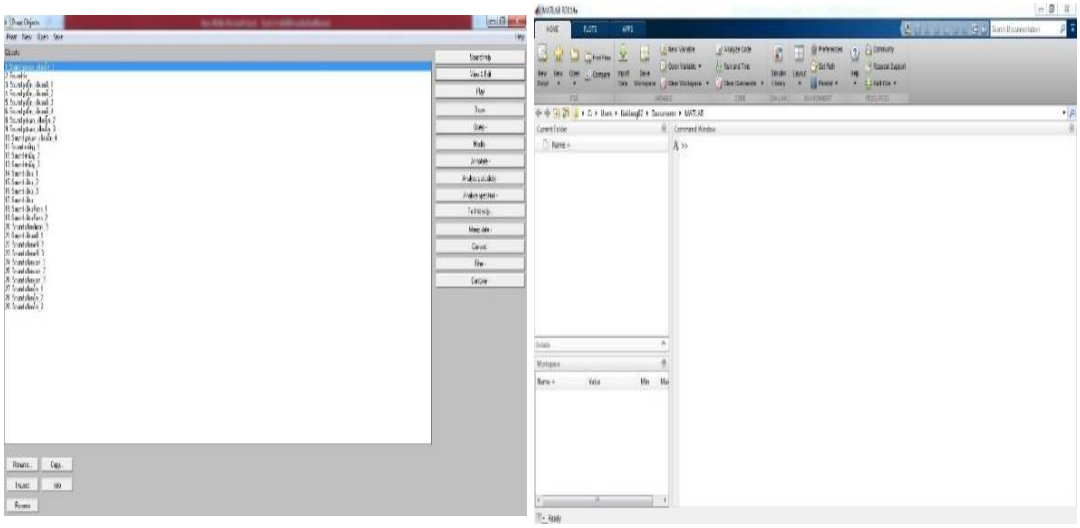


Figure 5: Praat and MATLAB programs.

2. Researchers took the data from the audio that has been converted and cut into words to analyze the techniques as follows: Spectrogram technique to consider the position of time, shape and size of the signal together with Wavelet Transform audio format conversion technique to convert the signal before comparing the similarity of the two data sets by using the Dynamic Time Warping (DTW) technique to measure the similarity between the sound signals.

3. Researchers took the converted and cut audio data and analyze by the Spectrogram technique to determine the position of time, shape, and size of the signal as shown in Figure 6.

UBRU International Journal Ubon Ratchathani Rajabhat University
Vol.3 No.3 September - December 2023

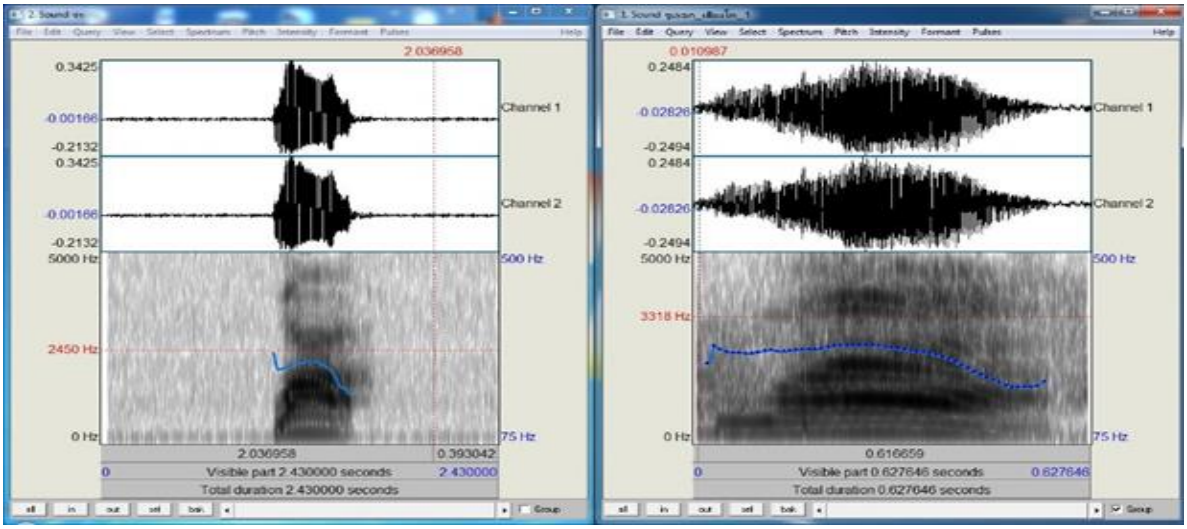


Figure 6: Spectrogram analysis.

4. Researchers took audio data that has been converted, and cut audio data, and analyzed with Wavelet Transform technique by compressing the original audio signal and sample to convert the existing signal accordingly before comparing the similarity of the data as shown in Figure 7.

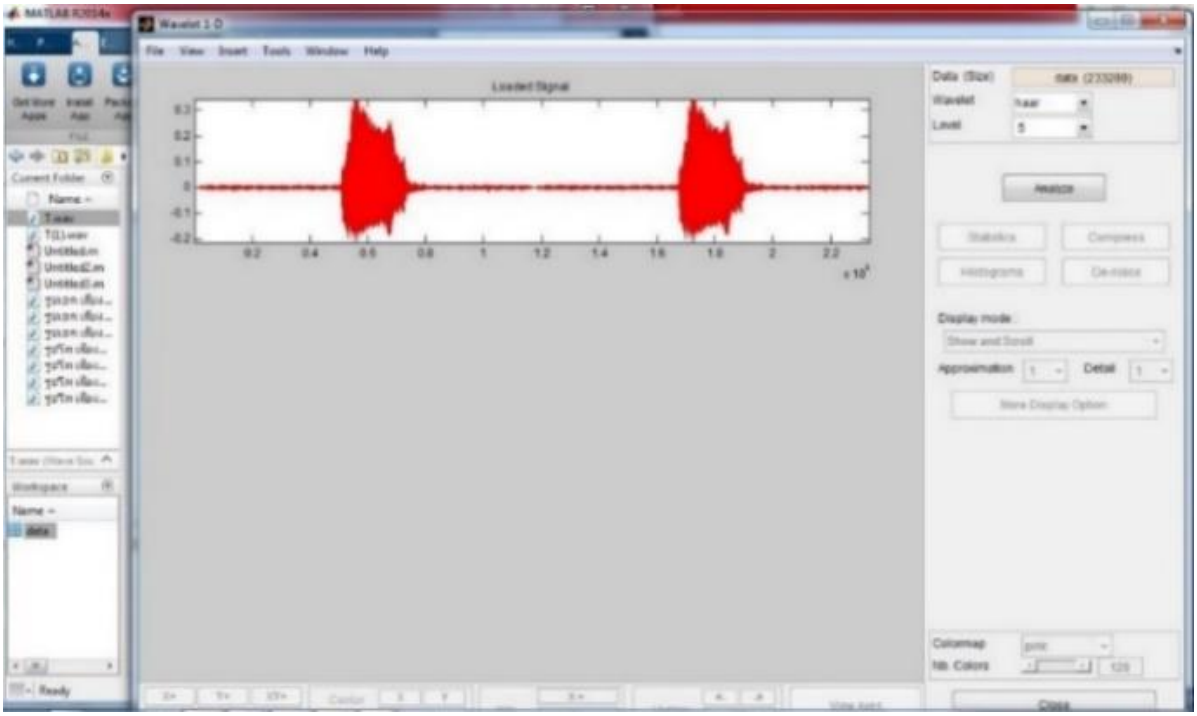


Figure 7: Analyze Wavelet Transform

5. Similarity analysis of 2 sound signal sets by using Dynamic Time Warping technique to find similarity of 2 data sets as shown in Figure 8.

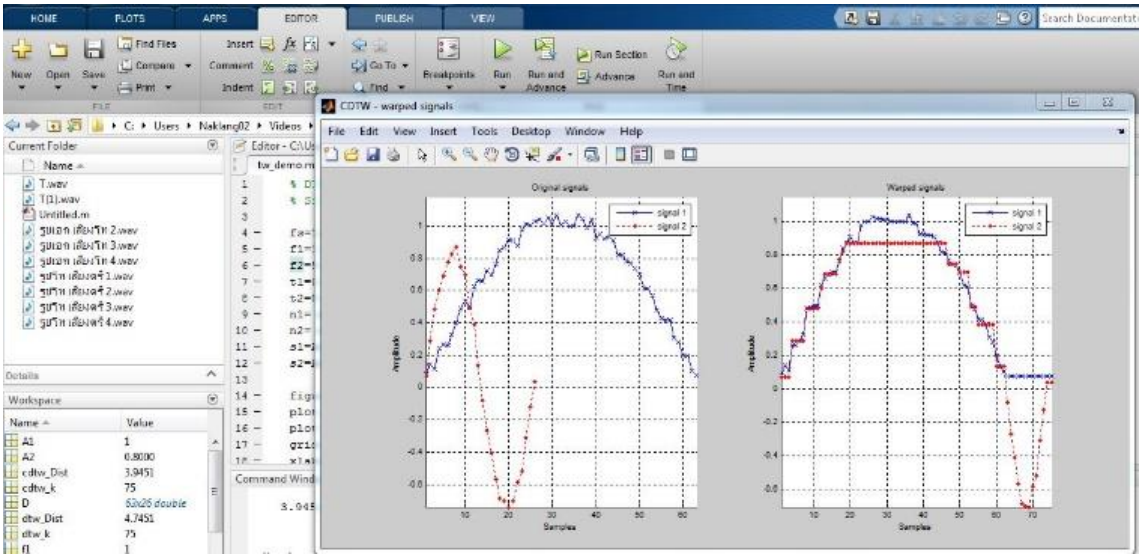


Figure 8: Similarity analysis of Dynamic Time Warping sound signal sets.

6. Deployment Phase

In this paper, a speech recognition system technique was used to help in detection of reading aloud and translating Thai Tones for young learners. Researchers used various techniques to analyze sound signal sets which could be applied with teaching other languages to help in problem solving with reading aloud and also serve as a guideline for developing a system or technology that helps in teaching language in the future.

Research Findings

Results of Data Analysis

Researcher presented the research results according to the following objectives:

Part 1 to study the techniques in data analyzing of speech recognition systems techniques to detect reading aloud skills in Thai tone indications or tonal conjugations for young learners

Table 4. Sound similarity results used by each technique together in the analysis.

Analytical Techniques			Sound Signal Resemblance (Percent)
Spectrogram	*Wavelet Transform*	*Dynamic Time Warping*	
✓			67.32
✓	✓		81.14
✓	✓	✓	93.84

From Table 4 shows the similarities of the sound signals used by each technique together. The first technique used was the sound wave spectrogram technique to determine the position of time, shape and size of the sound signal. 67.32 Percent, the next order will be the use of sound wave spectrogram technique together with Wavelet Transform sound format conversion technique to convert the existing sound properly to have a sound signal similarity equal to 81.14 Percent, the last order will be using the Wavelet Transform technique to optimize the existing signal before using the Dynamic Time Warping (DTW) technique to measure the similarity of two audio data sets. Together, each technique was able to compare and show that the sound similarity was 93.84 Percent higher.

Part 2 to study the effect of speech recognition systems techniques to detect reading aloud skills in Thai tone indications or tonal conjugations.

Table 5 Results of pronunciation training and comparison of sound signals

No.	**Results of pronunciation training and comparison of sound signals (number of times)**					**Average 5 times (Percent)**
	1	**2**	**3**	**4**	**5**	
1	86.21	88.22	87.12	89.03	90.50	88.22
2	90.39	90.93	92.68	93.79	93.82	92.32
3	91.87	92.79	93.76	93.82	93.84	93.22
4	86.48	88.87	87.62	89.94	90.97	88.78
5	87.95	89.97	90.59	91.42	93.14	90.61
6	88.91	89.17	90.29	91.52	93.74	90.72
7	88.98	89.95	91.98	92.80	92.88	91.31
8	85.98	86.45	89.01	91.21	92.42	89.01
9	85.98	88.87	87.96	90.94	91.98	89.15
10	92.95	93.22	93.79	93.82	93.84	93.52
11	86.98	88.06	89.91	91.95	92.91	89.96
12	90.92	92.23	92.98	93.79	93.88	92.76
13	86.91	89.91	90.30	91.52	93.74	90.48
14	86.98	88.16	89.91	91.92	92.95	89.98
15	90.95	93.22	93.79	93.82	93.84	93.12
16	85.56	85.82	87.19	89.18	91.02	87.75
17	86.91	89.91	91.38	92.82	93.84	90.97
18	85.98	88.66	89.41	91.89	92.88	89.76
19	84.98	87.66	89.21	91.82	92.41	89.22
20	87.50	89.85	90.88	91.82	93.84	90.78
21	89.57	90.85	91.89	92.98	93.82	91.82
22	85.50	86.62	87.19	90.23	92.71	88.45
23	86.98	88.66	89.21	91.92	92.82	89.92
24	90.97	91.97	92.98	93.82	93.84	92.72

No.	Results of pronunciation training and comparison of sound signals (number of times)					Average 5 times (Percent)
	1	2	3	4	5	
25	89.98	90.93	91.68	92.89	93.62	91.82
26	85.50	86.62	89.19	90.33	92.61	88.85
27	89.92	91.86	92.98	93.79	93.84	92.47
28	85.98	86.66	89.11	91.51	92.82	89.22
29	83.71	89.91	91.38	92.82	93.84	90.33
30	90.12	91.76	92.98	93.79	93.84	92.49
Total Average						**90.66**

From Table 5 the results of the pronunciation exercises and the analogy of the sound signals were compared. After using various techniques together to analyze 5 times, the average result is 93.52 Percent, the lowest mean is 87.75 Percent, and the total mean is 90.66 Percent**.**

Discussion

Discussions

1. The results of a study on the data analysis techniques of the speech recognition system to detect reading aloud and translating Thai tone indications or tonal conjugations for young learners found that the techniques involved in sound analysis consisted of Sound Wave Transformation technique, Spectrogram technique, Sound Wavelet Transform technique, and Dynamic Time Warping (DTW) similarity comparison technique by using the researcher have used various techniques together for analysis and comparison. A greater resemblance to an audio signal than using a single analysis technique. From the study of sound similarity by using various techniques together to analyze the similarity of the sound signal, the first technique used is the sound wave spectrogram technique to determine the position of time, shape and size of the signal to have a signal similarity. Votes accounted for 67.32 Percent. The next step will be to use the sound wave spectrogram technique together with the sound wavelet transform technique to convert the existing signal accordingly to have a sound similarity equal to 81.14 Percent. Sound wave spectrogram in combination with Wavelet Transform to optimize the existing signal before using Dynamic Time Warping (DTW) to measure the similarity of two audio data sets. By using each technique together, the analyzer can compare and show that the sound signal similarity is 93.84% higher. Translating sounds to homogeneous sounds for data analysis and is a method that has been effectively applied to speech recognition and signal processing (Korpunyapipat, 2015).

2. The results of using the speech recognition system to help practice reading skills in Thai tone indications or tonal conjugations for primary school students found that the results of practicing pronunciation and comparing the similarity of sound signals after using various techniques were analyzed 5 times. The highest average result was 93.52 Percent, the lowest average result was 87.75 Percent, the total average result was 90.66 Percent. The results of using the speech recognition system to detect reading aloud and translating Thai Tones for young learners found that the results of the pronunciation practice and the comparison of sound signals were similar. After using various techniques together to analyze 5 times, the average result is 93.52 Percent, the lowest mean is 87.75 Percent, the total average is 90.66 Percent,

consistent with teaching Thai language at present. with the use of technology and innovation in teaching and learning to help solve problems as well as to develop learners to learn effectively (Sinthaphanon, 2009). This result related to Malangpoo, Philuek, and Pomsamrit (2022) finding which showed that using speech recognition technology help teachers in detect the error of pronunciation in teaching Chinese speaking of Thai students

Conclusions

This research were to study the techniques in data analyzing of speech recognition systems techniques to detect reading aloud skills in Thai tone indications or tonal conjugations and to study the effect of speech recognition systems techniques to detect reading aloud skills in Thai tone indications or tonal conjugations, results shows that the technique in used for detect students' Thai Tonal Conjugation consisted of three techniques which involved Spectrogram Sound Wave Change Techniques, Wavelet Transform Audio Format Conversion Techniques, and Dynamic Time Warping (DTW). After experiment phase by using 30 students, it showed that the developed Speech Recognition Technology techniques can be used to detect students read aloud in Thai tone indications or tonal conjugations. Moreover, the next phase, researchers will develop the application which using this technique as the systems to detect how students read aloud and record in database.

Due to the limited sample size and other limitations of the lab setting experiment, these studies did not provide strong results that support the hypotheses that gestures help people comprehend and memorize information. Gestures were found beneficial for only one type of stimuli (mechanical systems) and an overall effect on memory test scores across text and diagram stimuli. Even though the effect of gestures was not significant between different types and formats of stimuli, it was in the right direction. Future research with more sensitive measurements could further explore gestures for thinking**.**

Suggestion

Suggestions for future research are as follows:

1. Future research should include additional variables that influence real earning management. These could include factors such as effective corporate governance, the type of audit firm, audit fees, and issues related to real earning management in relation to the operational performance of businesses, represented by financial ratios.

2. This study focused solely on companies listed on the Market for Alternative Investment (MAI). Therefore, future studies should expand the sample to include companies listed on the Stock Exchange of Thailand (SET), which are larger, have diverse investment forms, and a variety of shareholder structures.

References

Biometric Today. (2018). *5 Differences between Voice and Speech Recognition*. Retrieved from https://biometrictoday.com /5-differences-between-voice-and speech-recognition/.

Boonnithi, K. (2021). *Development of reading and spelling skills of first grade students using individual diagnostic exercises*. Independent Study of Master Thesis, Naresuan University, Phitsanulok, Thailand.

Cholkerkkait, C. (2018). *Development of Reading and Writing Skills by Using Skill Practice in Thai Language Core Content Area for Grade 1 Students*. Action Research, Wat Bomabring School, Surat Thani, Thailand.

Dipatee, K. and Narapongkasem, S. (2018). The Results of Thai Learning Activities Tonal Diversion and Tonal Writing by Using Thai Language Skill Practice Module for Prathomsuksa 3/3 Students at Nibongchanupathum School Mueang Disirict, Yala Province. *Proceedings of the 1st National Conference in Thai and Cultural Studies in Undergraduate Level,* Thaksin University, Song Kla, Thailand, pp 225-239.

Hanifa, R. M., Isa, K., Mohamad, S. (2017). Malay speech recognition for different ethnic speakers: an exploratory study. In: *2017 IEEE symposium on computer applications & industrial electronics* (ISCAIE)

Kershet, J., Bengio, S. (2009). *Introduction*. In: Margin, L., Methods, K., editors. Automatic speech and speaker recognition. West Sussex, United Kingdom: John Wiley & Sons Ltd, 2009.

Kikel, C. (2019). *Difference between voice recognition and speech recognition*. Retrieved from https://www.totalvoicetech.com/difference-between-voice-recognition-and-speech-recognition/.

Korpanyapipat, R. (2015). *Noise resistant recognition for Thai speech*. Master of Science, Thammasat University, Bangkok, Thailand.

Malangpoo, P, Philuek, W., and Pomsamrit, N. (2022) Using Speech Recognition System for Enhancing Chinese Pronunciation**.** *International Journal of Mechanical Engineering***,** 7(1), (January), 3442-3451.

Malaiwong, K. (1997). *Innovation in Technology 200 IT Perspectives*. (2nded.). Bangkok: Electronics and Computer Keyboard Technology Center, Bangkok, Thailand.

Mardyusoh, S. and Tornram, T. (2020). The development of learning achievement in Thai language at the elementary level 4th grad study by using tonal diversion exercises. Thamsuksamulniti School, Hatyai District, Songkhla Province, *Proceedings of Hatyai National and International Conference*, Songkhla, Thailand, pp 485-499.

Ministry of Education. (2020). *Core Curriculum Thai language learning subject group (Revised Edition B.E. 2560) according to the Core Curriculum of Basic Education BE 2551*. Retrieved on July 10, 2020, from http://academic.obec.go.th/ images/document/1559639270_d_1.pdf, 2017.

Na Nakhon, P. (1981). *Thai language problem*. (4thed.). Bangkok: Central printing house Office of the President King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.

Naiyaphat, O. (2005). *Quantitative and qualitative methods of analysis in behavioral sciences and social sciences*. (2nded.). Bangkok: Chulalongkorn University Printing House, Bangkok, Thailand.

Nasani, K. (2003). *Thai tones pronounced by tracheo-esophageal speakers: Phonemic analysis and perception testing*. In Thongkham, T. L. (editor), Thai accent of people without a voice box, pp. 97-118. Bangkok: Thailand Research Fund.

Panto, P. (2020). *Organizing teaching and learning in Thailand under the outbreak situation of coronavirus disease 2019 (COVID–19)*. Parliamentary radio station and academic offices, Bangkok, Thailand.

Phamontri, S. (2015). *The importance of the Thai language.* Retrieved from https://www.classstart.org/classes/4131.

Poosawad, S. (2018). *Comparison of Learning Achievement in Thai Tonal Conjugation of Grade 4 Students of Wat Don Yor School by Using Different Teaching Methods*. Action Research of Wat Don Yor School, Nakhon Nayok, Thailand.

Sharma, A. M. (2019). *Speaker recognition using machine learning techniques*. Master's Projects. Retrieved from https://scholarworks.sjsu.edu/etd_projects/685.

Sinthaphanon, S. (2009). *Innovations in teaching and learning to improve the quality of youth*. (3rd ed.). Bangkok: Chulalongkorn University Press, Bangkok, Thailand.

Suriyasophapan, C. (2003). *Basic frequency values of tones in the Thai language of people without a larynx and esophagus according to the duration of speaking training and the perception of tones by listeners*. In Teeraphan L. Thongkham (editor), Thai accents of people without a voice box, pp. 119-132. Bangkok: Thailand Research Fund.

The Secretariat of the Education Council. (2023). *Report on the results of the study of learning regression of basic education students in the COVID-19 situation: conditions, lessons learned, and guidelines for developing the quality of learning*. Retrieved from http://www.onec.go.th/th.php/book/BookView/1932

Thodthankun, T. and Chansuwan, P. (2016). The Study of Errors in Thai Writing: A Case Study of Students of Rajamangala University of Technology Suvarnabhumi Nonthaburi Campus. *Proceedings of The 1st RUSNC Rajamangala University of Technology Suvarnabhumi National Conference*, Nonthaburi, Thailand.

Thonglor, K. (1972). *Principles of the Thai Language Bangkok*. Bamrungsan, Bangkok, Thailand.

Watcharasukhum, S. (2013) *Problems in Using Thai Language from Thai - Malayu Students in Three Southern most Provinces*, Yala: Yala Rajabhat University, Thailand.

Zhi-Qiang, H., Jia-Qi, Z., Xin, W., Zi-Wei, L, and Yong, L. (2019). Improved algorithm of DTW in speech recognition. *IOP Conf. Series: Materials Science and Engineering*, 563, 052072. doi:10.1088/1757-899X/563/5/052072

Zsiga, E. and Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech*, *50*(3), 343–383.