OPEN ACCESS

Manuscript ID:
EDU-2025-13049251

Volume: 13

Issue: 4

Month: September

Year: 2025

P-ISSN: 2320-2653

E-ISSN: 2582-1334

Received: 30.06.2025

Accepted: 12.08.2025

Published Online: 01.09.2025

Citation:
Pomsamrit, N., & Waraaeksiri, S. (2025). A Study on the Appropriateness of Thai Words and Testing the Speech Recognition Technology Technique for Read-Aloud of Grade One Students in Thailand. *Shanlax International Journal of Education, 13*(4), 53–66.

DOI:
https://doi.org/10.34293/education.v13i4.9251

A Study on the Appropriateness of Thai Words and Testing the Speech Recognition Technology Technique for Read-Aloud of Grade One Students in Thailand

Nicharee Pomsamrit
Nakhon Sawan Rajabhat University, Thailand
https://orcid.org/0009-0000-9878-7996

Siwadol Waraaeksiri
Nakhon Sawan Rajabhat University, Thailand
https://orcid.org/0009-0002-7541-0058

Abstract
This research aimed 1) to collect and group Thai words and 2) to test the Speech Recognition Technology (SRT) technique for reading aloud. The research methodology was a mixed method which consisted of two stages belong to research aims: The first stage was a qualitative method to demonstrate the collection of Thai words for Grade One students by using a checklist table in 11 word categories to verify the details of Thai words and analyses by three Thai language experts, and the second stage was a quantitative method to test the SRT techniques in sound extraction by using the spectrogram sound feature extraction method and Mel Frequency Cepstral Coefficients (MFCC). The results showed that in stage 1, 2,423 Thai words were appropriate for Grade One students in11 categories; 1) single vowel-long sound, 2) single vowel-short sound, 3) spelled correctly according to the final consonant, 4) spelled incorrectly according to the final consonant, 5) tone marks, 6) no tone marks, 7) stable vowel, 8) transforming vowel, 9) reduced vowel, 10) words with diphthongs, and 11) words with leading consonants. stage 2, the ((MFCC) sound feature extraction method provided higher accuracy than the spectrogram. In future research, the researcher will apply the approach developed in this study to voice extraction, aiming to create an application that detects Thai reading aloud for Grade One students. This will enable testing of a wider variety of words and expand the target group in the future.

Keywords: Thai Words, Speech Recognition Technology, Sound Extraction, Mel Frequency Cepstral Coefficients (MFCC), Primary Education, Educational Technology

Introduction

The Thai Ministry of Education has decided to reform education by focusing on developing learners that cover knowledge, skills, and desired characteristics, according to the focus of learners in the 21st century. As shown in the Ministry of Education's educational policy announcement (The Royal Gazette, 2024), this requires the application of modern technology and artificial intelligence to help manage learning, especially promoting language learning under the principle of good knowledge and happiness, to produce visible results for learners. In particular, the reading ability of learners is considered an important basic ability for learners to learn in various subject groups according to the curriculum and in their daily lives. Therefore, the Ministry of Education and the Office of the Basic Education Commission have given importance to reading ability, especially at the first level of primary education (Grade 1-3). It has been determined as an urgent policy, the "Read, Write, and Do Math" Policy, which was created by the Ministry of Education. The National Institute

of Educational Testing Service (NIETS) is an agency involved in measuring and evaluating quality at both classroom and national levels.

Therefore, a project was established to assess the reading ability of grade1 students. The objectives were to enable Grade One students to understand their own ability to read aloud and understand. It also emphasises that schools should improve their teaching methods to solve this problem by distributing spelling words, promoting the use of local languages, and Thai as a medium for teaching diverse languages, to lay the foundation for students to develop their analytical thinking skills, as well as their communication skills and the use of a third language to extend their learning effectively.

There are Thai language scholars who have given their opinions about Thai pronunciation, as Satyophas (2002) stated that reading aloud is a continuous process between the eyes, brain, and pronunciation. That is, the eyes will focus on the letters, and the brain must translate the letters into thoughts and then process the thoughts into words, along with the pronunciation out loud. Sathira-angkool (2003) stated that pronunciation is the correct pronunciation of letters in Thai, which has many criteria and is different. However, it must be correctly read according to the alphabet. Kaewsanae (1999) stated that pronunciation is the correct pronunciation of the letters that appear according to the alphabet or language by using the organs related to pronunciation, such as the lips, roof of the tongue, teeth, or vocal cords, to move according to the characteristics of the base where the sound source is produced. According to Churchun (1990), there are many types of pronunciation, stating that the general principles of reading aloud are as follows:1) Correctness means that the reader can read aloud correctly according to Thai or other language spellings that Thais uses, including reading correctly according to popularity. The reader must study the principles of correct reading from books, observe and collect words, and correct pronunciations consistently, 2) Clarity means pronouncing words clearly, including vowels, consonants, tones, and diphthongs. They must be read clearly and audibly and not too loudly or soft. 3) Fluency refers to being able to read aloud continuously without interruption or rhythm loss. This fluency comes from frequent and consistent practice until reading skills are developed, which will lead to fluency in reading, and 4) pausing time or punctuation is important in reading aloud because if the reader pauses time or punctuation incorrectly, the listener may misunderstand the meaning.

In terms of the problem of pronunciation among Thai language learners, scholars have proposed the following ideas about the problem: Ubonyam (1999) mentioned the problem of reading aloud that it is necessary to adhere to the principle of understanding what is to be read, such as understanding every word, every passage, and correct pronunciation, which is consistent with Changkhwanyuen et al. (2006) and Premphan (1999), whostated that reading aloud must pronounce words correctly, clearly, accurately, fluently, and must be practiced and well prepared, as well as who stated that one must read clearly, including pronouncing consonants, vowels, and tones correctly, such as pronouncing consonants that require the tongue to be rolled or diphthongs.

This reading problem is evidenced by the exam results in past 5 years show the significance of the problem of “Words with leading consonants”, which has the lowest score from the test for 5 consecutive years (from more than 600,000 students taking the test each year) as follow: 1) in 2024 Reading Test results found that the lowest achievement clusters were “Words with leading consonants”, followed by Words with diphthongs, 2) in 2023 Reading Test results found that the lowest achievement clusters were “Words with leading consonants”, followed by Spelled incorrectly according to the final consonant, 3) in 2022 Reading Test results found that the lowest achievement clusters were “Words with leading consonants”, followed by Spelled incorrectly according to the final consonant, and Tone marks, 4) in 2021 Reading Test results found that the lowest achievement clusters were “Words with leading consonants”, followed by Tone marks, and 5) in 2020 Reading Test results found that the lowest achievement clusters were “Words with leading consonants”, followed by Tone marks (Office of the Basic Education Commission, 2025), which related to Sumrongpan & Erawan (2022) study who claimed that there were 7 problems were found in

learning Thai language of Grade One students: 1) Consonants, 2) Vowels, 3) Intonation marks, 4) Words with no final consonant, 5) Words with final consonants, 6) Words with diphthongs, and 7) Words with leading consonants. Therefore, most research focuses on teaching and learning processes, such as Waibanthao's (2020) Learning management by using tales to develop the Thai reading abilities of Grade 1 students. Sukkeaw (2022) compared the reading abilities of students with pre-test and post-test on learning with Brain-Based Learning of Grade One students, which gave post-test results higher than pre-test. Hence, this research focuses on developing a tool to help check students' speaking skills, which is an innovation in teaching Thai today. In Thailand, the use of technology to measure and evaluate language learning outcomes is not as widespread as it should be, and the lack of technology in the measurement and evaluation of Thai language learning is a significant issue that hinders effective educational practices. Despite the potential benefits of integrating technology, various barriers prevent its widespread adoption in language assessments. Numerous classrooms in Thailand continue to depend on conventional teaching approaches with minimal incorporation of technology for evaluating language skills. Frequent challenges include inconsistent Internet access, inadequate technological tools, and a shortage of instructional strategies (Teemuangsai and Meesook 2017). However, Warden and Ye (2017) highlighted that although students strongly anticipate the use of ICT, the reality of its implementation often does not meet these expectations, especially in language learning environments. Moreover, the lack of automated pronunciation evaluation tools, such as those implemented in English language learning, indicates that comparable technologies might improve Thai language assessments. Therefore, this study presents the development of a Thai pronunciation measurement system for children, showing the developmental steps from studying Thai words to displaying sound analysis techniques to be used as a guideline for further development of a complete application.

Related Literature

Thai Language Words

Educators and various organisations present knowledge about the Thai language, the most widely known among Thai people is Thai Language Principles: Spelling, Verbology, Syntax, and Prosody, written by Phraya-Upakitsanlapasarn (1995). This is an explanation of the Thai language according to traditional grammar. This textbook has long been used in the field of linguistics. Considering the content, it can be said that, although there are some questionable analyses, there are many places that are profound and logical and have been used as a basis for syntactic analysis. Other textbooks in the same vein as Phraya-Upakitsanlapasarn (1995) with the same name include Thai Language Principles by Thonglor (1982), Thai Language Principles by Phongphaiboon (1988), and Thai Language Principles by Chanawong (1986). It can be said that all three books have the same concept as Phraya-Upakitsanlapasarn (1995) in every respect. Other books in the same vein, with some differences, include Thai Grammar by Lamduan (1983), Siam Grammar, Prosody, and Thai Language Textbooks by the Academic Division. The Ministry of Education (1937), one of the Thai grammar content that students must learn, includes Thai words.

The Royal Institute Dictionary (Royal Institute 2011) states that a sound uttered at a time, a spoken sound, or written or printed text to express an idea is usually considered the smallest unit that has meaning in itself, used in front of other words to have meaning, while the Faculty of Thai Language, defines a word as a sound uttered or a combination of letters that have the meaning of some kind; each word is different in both function and meaning in language use. Therefore, users must know the types of words and their meanings very well to be able to select words for effective communication. Learning about the types of Thai words allows learners to use words correctly and appropriately for each person; therefore, the types of Thai words have been defined.

Changkhwanyuen et al. (2006) define a word as the smallest unit of language that native speakers know and use in speaking and writing. A word is a language unit that can be used alone for communication purposes. Learning words is the

foundation for learning the entire language system, and in conclusion, a word refers toa sound uttered by speaking or written text. With its meaning to communicate with each other as an important component in sentences, words have different meanings; the meaning of words depends on their function.

Anchalinukul (2003) provides an overview of the study of words in Thai in all aspects. Because it mentioned the concept of classifying the types of words in Thai according to various criteria, including the classification of words according to word formation, sound level or word level, the classification of words according to traditional grammar and structural grammar, the division of words according to their function in sentences and the division of words according to their meaning.

In learning Thai words, there are studies showing problems with learning Thai words as follows: the results of the report of the National Test of Basic Ability of Students (National Test: NT) at the 3rd grade of Nong Rangsit School from the academic year 2018-2020, the scores in the language ability in the indicator P.3/1, spelling words, and telling the meaning of words decreased. In the academic year 2020, the average score was 32.14 (Nong School, 2021), a score that should be urgently improved. The problem is that students lack the skills to practice writing and spelling words correctly;therefore, students must understand the language and spelling rules better. Words that are not spelled according to the spelling rules, especially the mother syllables, Mae Kok, Mae Kot, Mae Kon, and Mae Kob, are words with many spellings that do not conform to the rules, causing students to be confused when writing the words. It can be said that writing the correct spelling will allow students to convey meaning well and correctly, according to their intentions of communication.

As Laimoon and Aroonmanakun (2017) states, writing is an important method of communication. This is one method for humans to communicate with each other. It is also a method used to convey knowledge, ideas, and experiences between the communicator and the receiver. Therefore, writers must have the skill to write and spell words correctly.

Speech Recognition Technology

Speech Recognition Technology is now increasingly used in pronunciation studies and speech training because it is more accurate in detecting sounds at the frequency level than the human ear and has higher standards. Speech recognition is the process of extracting a word sequence from an audio signal. When Bell Laboratories created the “Audrey system” in 1952, voice-speech recognition was first recorded. Later, IBM created “Shoebox”, which was credited with being the precursor to modern voice recognition systems and was able to identify and respond to sixteen English words. In 2011, Apple introduced Siri, a voice recognition technology that has steadily gained popularity and gained widespread acceptance thanks to Google Voice Search and Amazon Alexa (Vaidya et al. 2021)

Automatic speech recognition, the process by which a computer transforms an acoustic voice signal into text, is the use of speed recognition technology. The method by which a computer converts an acoustic speech signal into an abstract interpretation of speech is referred to as automatic speech understanding, which includes several approaches to Speech Recognition Technology, such as Acoustic-Phonetic Pattern Recognition Artificial Intelligence (Anjali and Sherseena, 2020).

In addition, researchers such as Ahuja (2025) have presented a comprehensive modern ASR technology with challenging techniques and applications. Also, Malangpoo et al. (2022) studied the use of speech recognition technology to detect Chinese pronunciation in Thai high school students and found that it can help students practice pronunciation and receive immediate feedback. Philuek and Puttasem (2023) studied three techniques that involve sound: Spectrogram Sound Wave Change Techniques, Wavelet Transform Audio Format Conversion Techniques, and two sets of data similarity comparison techniques called Dynamic Time Warping (DTW). Moreover, Passon and Philuek (2024) developed an application to detect read aloud based on Thai words in the reading test lists which showed that there was an accuracy of more than 50%, which was 93.75%, and less than 50%, which was 6.25% after using the application to detect pronunciation which means that the accuracy was at the highest level.

Sound Extraction: Mel Frequency Cepstral Coefficients (MFCC)

Mel Frequency Cepstral Coefficients (MFCC) are one of the numerous methods that have been modified as sound extraction Technique to incorporate Speech Recognition Technology into machines and systems from the beginning of the development of speech recognition systems. Assuming that the human voice is a good speaker recogniser that is dependent on the vocabulary of reasonable-to-high size, the MFCC is a replication of the human hearing mechanism that selectively applies the operating theory of the ear. Therefore,it is frequently used to identify and comprehend speech (Ali et al., 2021).

MFCC is a widely used framework feature extraction method that is often cited as beneficial for a number of applications. The goal of speech signal analysis is to uncover more pertinent, concise, and instructive information than raw data of the speech signal. One of the most well-known representations of speech analysis is vocal tract features, commonly referred to as segmental, spectral, or systemic features (Ramakrishnan, 2012).

The MFCC has been used with a variety of classifiers, including those that are appropriate for the global representation of features (such as SVM, kNN, ANN, and MLC) and those that are used with time-series representation (such as DTW, HMM, and LSTM), according to an analysis of 187 publications. According to the review, the program determines the classifier with the MFCC feature that is most frequently utilised (Abdul and Al-Talabani, 2022).

Methodology

The purpose of this research was to study the appropriateness of Thai words and test the Speech Recognition Technology technique for reading aloud of Grade One students in Thailand, using mixed methods research divided into two stages: first, qualitative research was used to synthesise Thai words and classify suitable words for Grade One students, and second, quantitative research was used to present the development process of a Speech Recognition Technology system for testing the pronunciation of words.

Stage 1 was qualitative research to synthesise Thai words and classify suitable words for Grade One students.

Instruments used in this study. The instrument for analysing Thai words appropriate for Grade One students is a checklist for checking the details of Thai words for Grade One students. Thai words were obtained from the Report on the Study of Basic Vocabulary of Primary School Students, Grades 1-3, Academic Year 2011 (Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission, 2011).

Data Collection

Qualitative data were collected using a detailed checklist by categorising the words into groups according to word type. The researchers invited three Thai Language experts (the experts consisted of two university lecturers and one primary school Thai language teacher) to a meeting to organise words into 11 groups, which were divided into 11 groups of words as follows: single vowel-long sound, single vowel-short sound, spelled correctly according to the final consonant,s pelled incorrectly according to the final consonant, tone marks, not one marks,s table vowels, transforming vowels, reduced vowels, words with diphthongs, and words with leading consonants. The Report on the Study of Basic Vocabulary of Primary School Students, Grades 1-3, Academic Year 2011 and 2025 (Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission, 2011, 2025) found that 1,007 Thai words were ranked in terms of reading difficulty, and 993 words (Year 2011 list) and 708 words (Year 2025 list) were selected for use in spoken language for Grade One students (examples of words are shown in the table in Figure 1).

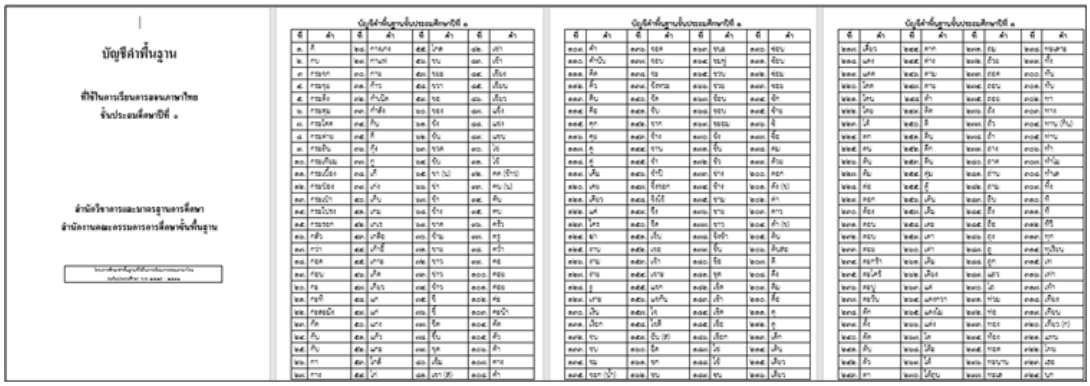

Figure 1 Examples of Thai Words from the Report of Bureau of Academic Affairs and Educational Standards (Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission, 2025)

Analysis

The researchers conducted a meeting to organise the words into groups and eliminate words that were not words that Grade One students should learn (words that were too difficult and inappropriate for the level of the students). Each of the 613 words could be classified into several 11-word categories. The results of the study found that when organized into 11 word categories, a total of 2,423 words could be displayed (as shown in Table 1 in the results of the study), which the researchers could use in further studies.

Stage 2 was quantitative research to present the development process of a Speech Recognition Technology system for testing the pronunciation of words

Instruments used in the research. Tools/Systems used in developing the speech recognition technology system consisted of a computer,and techniques used in data analysis were spectrogram mel-frequency cepstral coefficient (MFCC) techniques, together with the program for database, sound record, and edit files such as Ocenaudio program, Visual Studio Code program, to design a database using PHP Myadmin and FastApi.

Data Collection

The researchers proceeded by selecting 20 Thai words and having 20 children pronounce each word three times for use as a preliminary test of the analysis technique to see if the phonological analysis process could be further developed.

Data Analysis

The research process was adapted from the CRISP-Data Mining model, with each step of the study as follows:

Step 1: Understanding the Problem

Researchers have studied related research and information about the listening problems of reading test examiners, Speech Recognition Technology, using the spectrogram technique, and the Mel Frequency Cepstral Coefficient (MFCC) technique to analyse the sound patterns by training data that will lead to the system and tools.

Step 2: Understanding the Data

The researchers studied the existing data and considered the possibility of analysing them. In this step, the researcher selected children who could read aloud to record the sounds from the Thai words we chose, 20 words from 20 students who could read, totalling 1,200 sounds collected to be used as prototype sounds for sound analysis.

Step 3: Data Preparation

The process of preparing data by converting the raw data to the data to be used in the next step includes selecting the variables to be analysed appropriately and converting the format of the variables to the same format to prepare the data for modelling. The data were divided into three substeps.

From the sound data collection, the researcher converted the obtained sound data, cut the sound into words, changed the sound into Mono format, and converted the file to a.wav file for further analysis.

The researchers used the Ocenaudio program with the following methods:

- Import sound data to use the Ocenaudio program;
- Convert the file from Stereo to Mono, making the file size small and easy to calculate using the Ocenaudio program
- Cut the sound into each word using the Ocenaudio program
- Cut the sound into each word by recording the sound as a.wav file and storing the data in a folder using the Ocenaudio program.

From cutting the sound, divide the data into 20 words and adjust the sound format, time, frequency, and values used in the calculation to match the data in each word. To easily analyse the data, the researchers have a method of writing code to adjust the time and frequency values used in the calculation to be the same, as shown in Figure 1. The code used to adjust the time and frequency values was the same.

Step 4: Data analysis

This is to take audio data that have been converted from file formats and cut into words, and use it to analyse and compare the similarity of the sounds with the following steps:

The researchers chose to use the following analysis techniques:

- Spectrogram technique and
- Mel Frequency Cepstral Coefficient (MFCC) technique toanalysethe data using the Jupyter Notebook tool to create a Python Machine Learning model (Figure 2).

Figure 2 Python Machine Learning Model by Jupyter Notebook

Create a learning model using the TensorFlow and Keras libraries. To create the Deep Neural Network, 20 words of audio data were used. Each word used the spectrogram technique and Mel Frequency Cepstral Coefficient (MFCC) technique.

Test the sound with 1 initial sound and all 20 trained sounds to obtain the accuracy and error values, such as the value of the Spectrogram sound feature extraction method and the value of the Mel Frequency Cepstral Coefficients (MFCC) sound feature extraction method.

Results

The researchresults are presented in two parts:

- Part 1: Results of the analysis of words suitable for Grade One students
- Part 2: Results of the test to present the process of developing a speech recognition system to further develop applications, as follows:

Part 1: Results of the Analysis of Words Suitable for Grade One students

NO.	Single vowel-long sound	Single vowel-short sound	Spelled correctly according to the final consonant	Spelled incorrectly according to the final consonant	Tone marks	No tone marks	Stable vowel	Transforming vowel	Reduced vowel	Words with diphthongs	Words with leading consonants	Total
1	-	ค่ะ	-	-	ค่ะ	-	ค่ะ	-	-	-	-	3
2	-	ครับ	ครับ	-	-	ครับ	-	ครับ	-	ครับ	-	5
3	-	ไม่	-	-	ไม่	-	-	ไม่	-	-	-	3
4	-	กี่	-	-	กี่	-	-	-	กี่	-	-	3
5	มี	-	-	-	-	มี	มี	-	-	-	-	3
6	หนู	-	-	-	-	หนู	หนู	-	-	-	หนู	4
7	แล้ว	-	แล้ว	-	แล้ว	-	แล้ว	-	-	-	-	4
8	-	แล้น	แล้น	-	แล้น	-	แล้น	-	-	-	-	4
9	-	ใส่	-	-	ใส่	-	ใส่	-	-	-	-	3
10	ที่	-	-	-	-	ที่	ที่	-	-	-	-	3
11	-	เป็น	เป็น	-	-	เป็น	-	เป็น	-	-	-	4
12	ขอบ	-	ขอบ	-	-	ขอบ	ขอบ	-	-	-	-	4
13	เคย	-	เคย	-	-	เคย	-	-	เคย	-	-	4
14	อยู่	-	-	-	อยู่	-	อยู่	-	-	-	อยู่	4
15	แม่	-	-	-	แม่	-	แม่	-	-	-	-	3
16	ชื่อ	-	-	-	ชื่อ	-	ชื่อ	-	-	-	-	3
17	-	ไป	-	-	-	ไป	-	ไป	-	-	-	3
18	มา	-	-	-	-	มา	มา	-	-	-	-	3
19	-	ให้	-	-	ให้	-	ให้	-	-	-	-	3
20	ว่า	-	-	-	ว่า	-	ว่า	-	-	-	-	3

Figure 3 A Table with the Frequency of Each Word Belonging to 11 Categories

Figure 3 shows an example table that was used to calculate the frequency of Thai words in each category. This image is only a part of the entire table. When all 613 words are entered into the table to count their frequency, the details are as shown in Table 1, which summarises the number of words and examples of Thai words.

Table 1 The Example of Some Thai Words for Grade One Students in 11 Categories

Word Categories	Frequency of Words	Thai Words (Example in Thai Language)
Single vowel-long sound	369	ดวงต่างผ่านย้ายแรงลองดื่มเดิมถ้วยถีบแถมหนูร้อนเหลือ
Single vowel-short sound	244	ดังนุ่นป้อมผงพริกยอมครับเลขคุณอาจสุขใหม่เหมือนกันใกล้
Spelled correctly according to the final consonant	429	ดังดื่มเดิมถ้วยถีบแถมนุ่นปวดป้อมผงพริกแล้วร้อนเหมือนกัน
Spelled incorrectly according to the final consonant	55	รูปรถเลขคุณบาทศาลอาจญาติบอลเสร็จสุขภาพเขตแหละ*
Tone marks	219	ดื่มนุ่นป้อมผงแล้วน้ำว่ายบ่งป่วยเขียวใหม่ร้อนใกล้ตั้ง
No tone marks	397	ดังเดิมถ้วยถีบแถมปวดแผลพริก มวลยอมโยนหนูเขตแหละ*
Stable vowel	475	ดื่มถีบแถมนุ่นป้อม แผลพริกภูยอมโยนหนูแล้วอยู่เขตแหละ*
Transforming vowel	77	ดังเดิมปวดเกิดกลับ เสร็จรัก ฝน ลัก ตันตัด เปิด ดับ ตัด
Reduced vowel	66	ถ้วยผง มวลลดกี็รถชน บวก ล้ม กัด ผล รส อ้วน กลบ ครบ
Words with diphthongs	55	ครับครัวกราบแกล้งแกว่งคล่องสระหมุนกรีดกลมครบ
Words with leading consonants	37	หนูอยู่หรอกหลับเหงาหน่อย ใหม่ ใหม่เหมือนเหลือหลังแหละ*
Total	2,423	

Table 1 shows that the total number of words used was 2, 423. It was found that the top three vocabulary items for Grade One students were stable vowels (475 words), followed by spelled correctly according to the final consonant (429 words), and not one marks (397 words). In comparison, the top three clusters were Words with leading consonants (37 words), Spelled incorrectly according to the final consonant, and Words with diphthongs (55 words), respectively. In addition, some examples of words can belong to multiple word groups, such as"แหละ",which belongs to No tone marks, Stable vowel, Spelled incorrectly according to the final consonant, and Words with leading consonants categories.

Part 2: Results of the test to present the process of developing a speech recognition system to further develop the applications.

According to Figures 4–7 in Figures 4. Error value of the spectrogram sound feature extraction method. Figure 5. Accuracy value of spectrogram sound feature extraction method. Figure 66. Error value of the Mel Frequency Cepstral Coefficients (MFCC) sound feature extraction method. Figure 7. The accuracy value of the Mel Frequency Cepstral Coefficients (MFCC) sound feature extraction method is to be used in selecting the method of extracting sound pattern analysis features over the use of the Spectrogram technique

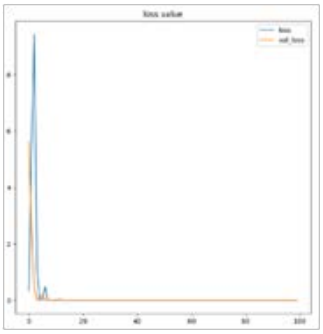

Figure 4 Error Value of the Spectrogram Sound Feature Extraction Method

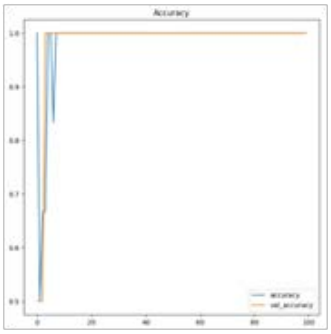

Figure 5 Accuracy Value of the Spectrogram Sound Feature Extraction Method

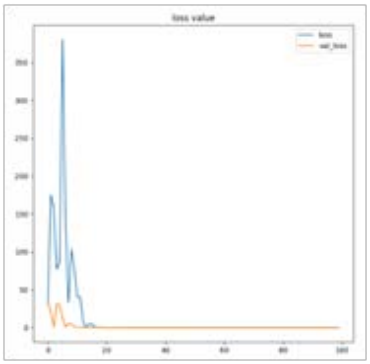

Figure 6 Error value of the Mel Frequency Cepstral Coefficients (MFCC)

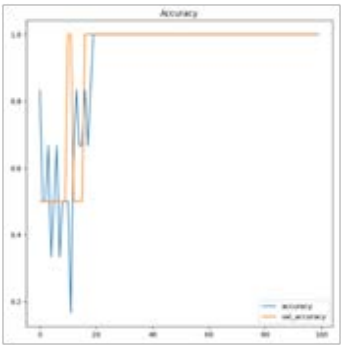

Figure 7 The Accuracy Value of the Mel Frequency Cepstral Coefficients (MFCC)

Table 2 Accuracy and Error Values for Selecting the SRT Methods Analysis Feature Extraction Method

Words	Spectrogram		Mel Frequency Cepstral Coefficients	
	Accuracy	Error	Accuracy	Error
ผง (Phong)	1.000	4.768	1.000	0.005
มวล (Moun)	0.879	3.244	1.000	0.006
ลด (Lot)	1.000	2.335	1.000	0.003
ก็ (Kor)	0.956	3.344	1.000	0.007
รถ (Rod)	0.885	0.921	0.966	0.002
ชน (Chon)	0.981	0.018	1.000	0.007

After testing sounds with 1 initial sound and 20 trained sounds in each word, Table 2 shows the results of using 2 techniques in testing, and demonstrates the number of sounds in each word with the accuracy and error values.

Table 3

Student NO./Vocab	ผง (Phong)	มวล (Moun)	ลด (Lot)	ก็ (Kor)	รถ (Rod)	ชน (Chon)
Student NO. 1	0.945	0.953	0.945	0.961	0.984	0.977
Student NO. 2	0.971	0.971	0.956	0.954	0.964	0.964
Student NO. 3	0.985*	0.973	0.973	0.941	0.988*	0.986*
Student NO. 4	0.964	0.952	0.951	0.945	0.976	0.973
Student NO. 5	0.972	0.971	0.953	0.944	0.985	0.975
Student NO. 6	0.977	0.977*	0.961*	0.963*	0.986	0.971
Total	0.969	0.966	0.956	0.951	0.980	0.974

After using the sound analysis technique and showing the results as in Table 2, the researcher tried using the sound analysis from the experiment with the Grade One student, which is shown in Table 3, which shows the sound frequency value (1 is the original sound) from the sample of pronouncing 6 Thai words from the sound analysis sample of 6 students to conclude that in the overall picture by analyzing the average value, it was found that,

- The first word (ผง (Phong)), Student No. 3, was able to pronounce the most clearly (0.985*), and the overall pronunciation,withan average value of 0.969.
- The second word (มวล (Moun)), Student No. 6 was able to pronounce the most clearly (0.977*), and in the overall pronunciation, it was found that the average value was 0.966.

- The third word (ลด (Lot)), Student No. 6, was able to pronounce the word most clearly (0.961*), and in the overall pronunciation, it was found that the average value was 0.956.
- The fourth word (กี๋ (Kor)), Student No. 6, was able to pronounce the most clearly (0.963*), and in the overall pronunciation, it was found that the average value was 0.951.
- The fifth word (รถ (Rod)), Student No. 3, was able to pronounce the most clearly (0.988*), and in the overall pronunciation, it was found that the average value was 0.980.
- The sixth word (ชน (Chon)), Student No. 3, was able to pronounce the most clearly (0.986*), and in the overall pronunciation, it was found that the average value was 0.974.

Conclusions

The research results showed that,

There were 2,423 words Thai words that appropriated for Grade One students in 11 categories; 1) Single vowel-long sound, 2) Single vowel-short sound, 3) Spelled correctly according to the final consonant, 4) Spelled incorrectly according to the final consonant, 5) Tone marks, 6) No tone marks, 7) Stable vowel, 8) Transforming vowel, 9) Reduced vowel, 10) Words with diphthongs, and 11) Words with leading consonants which can be used to select vocabulary by word group for further development of application models to evaluate Thai word pronunciation. The results of this research in word analysis were conducted by qualitative research by analyzing the vocabulary of each word and putting it into a word group analysis table. The vocabulary used was standard words that Grade One students must know according to the report of Bureau of Academic Affairs and Educational Standards (Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission, 2025).

The ((MFCC) sound feature extraction method gives higher accuracy than the spectrogram, as shown in Table 2, and the results are shown in Table 2. Students who lack pronunciation skills will find it difficult to predict the sounds of unknown words, which affects reading at a higher level. The events that indicate that students have reading problems are published in the research of Lanchwathanakorn (2016) as follows: 1) students cannot combine words correctly, such as "รถ (Rod)" (Ro-O-Dor = ROD). Students cannot tell how many syllables there are in the word they read and have problems spelling rhyming words or spelling new words from consonants or vowels, and 2) students cannot decode the meaning and sound of words, including students having difficulty matching sounds and letters, and students having difficulty reading and spelling to pronounce words. In addition, research by Phongprasert and Makjui (2021) and Kuhapinan (1999) has presented common pronunciation problems, which are divided into two types: dyslexia and misreading.

Speech Recognition is a technology that helps create interactions between humans and computers through natural speech. Human speech is a highly variable type of data. For example, if 3 people say the same word "I", the first person is a man, so his speech frequency is low. The second person was a woman, so his speech frequency was high. The third person is a foreigner; therefore,his accent is different from that of a Thai person. Several factors affect speech. Even if the same word is said, the sound waves emitted may not be the same. One of the evaluations of the quality of Speech Recognition Technology can be performed by having AI receive speech data and translate the speech into text. The text generated by AI was then compared with the correct text. Then, we determine the word error rate (WER) in the text. Researchers have compared the Speech Recognition models of leading companies that provide speech recognition services, including Google, Microsoft, and Amazon, with those developed by Thais. All models compared were Google, Microsoft, Amazon, NECTEC, and AI. The WER of each model was 13.711, 12.579, 21.863, 20.106, and 13.673 %, respectively. This shows that the current Thai speech recognition can work well, but there are still some errors (Suchato et al. 2023).

The MFCC has been used with a variety of classifiers, including those that are appropriate for the global representation of features (such as SVM, kNN, ANN, and MLC) and those that are used with time-series representation (such as DTW, HMM, and LSTM), according to an analysis of 186 publications. The majority of studies have accepted the usage of

a worldwide representation of the MFCC, which is primarily globalised by computing the statistics and features along the frames, despite the fact that the MFCC is computed in a short-time signal. Further research is necessary to demonstrate the potential of time series versus non-time series in modelling a range of applications (Abdul and Al-Talabani, 2022). In this research, it was found that the Mel Frequency Cepstral Coefficients (MFCC) technique gave the highest accuracy value, suitable for use in analyzing Thai word sounds. The research results are consistent with the research results that found that MFCC is suitable for analyzing human speech signals (das et al. 2014; Passon & Philuek, 2024 and Philuek & Puttasem, 2023) or used as an extraction technique (Muda et al. 2010). The implementation of Speech Recognition Technology for reading aloud can be achieved by having AI assess the accuracy of reading for students by having them say the desired word, and then having AI decode the words that the children read. If the decoding matches the word that we asked the children to read, it means that the children read it correctly. If it does not match, this means that the children read it incorrectly.

However, scoring reading by Speech Recognition can only roughly assess ability because training the Speech Recognition model focuses on having the model predict words correctly. Therefore, if a student pronounces incorrectly, such as trying to say the words "Wang-Nee-Peng-Eang-Song-Kang" in the Google Cloud Speech-to-Text program, the model will decode the words Today as Songkran Day. In this case, the model will not be able to detect the students' mispronunciations. To create a model to assess students' reading ability, it may be necessary to change the model from Speech Recognition to Phoneme Recognition instead, or translate speech sounds into phonemes, which are the smallest units in a language, as in the Elsa Speak application. However, at present, there is no Thai language dataset; therefore, it is not possible to create a model (Suchato et al. 2023).

Limitation

This paper presents the results of a qualitative study of Thai word analysis and a quantitative study of sound analysis techniques using sound wave analysis. This study aims to develop a system/ application for measuring pronunciation, which is expected to be more accurate than having teachers sit and listen toaudio and provide scores. This is a pioneering study onthe use of Speech Recognition Technology in Thai language studies, and therefore may have limitations and recommendations for further research, as follows:

Because this is new research with limitations regarding the development of Speech Recognition Technology for assessing Thai reading sounds for Grade One students, it is necessary to analyse words appropriate for learners at this level. Furthermore, it must be consistent with the reading tests administered by the Office of the National Education Commission, Ministry of Education, based on past test results that show scores for each category of words. Therefore, this research demonstrates the 11 categories of words, and further studies will focus on selecting words appropriate for application development.

This study selected only two sound analysis techniques, and in future research, it is necessary to add other sound analysis techniques to improve the accuracy of the system.

This study used a sample of 20 children, selecting 20 words,with each child saying each word only three times. The test results showed that the MFCC technique produced the fewest errors. However, the researchers had limitations in processing; therefore, a wider variety of sounds and words should be added in future studies.

Further Study

Application Development for the Future Study

Further research is needed to develop an application using Speech Recognition Technology to classify the results of Thai pronunciation according to the reading test of Grade One students. Researchers have used various techniques to analyse the sound signal set along with the application usage pattern, which can be applied to Thai language teaching to help solve the problem of learners' pronunciation, and also serve as a guideline for developing systems or technologies that help in teaching language in the future.

In this section, application development guidelines are presented to provide a visualisation of

its utilisation and future research, as detailed below.

1. The first step in developing a Speech Recognition Technology application to classify Thai words read a loud is to enquire about the needs of the system users. The researchers had to ask the Thai language teachers of Grade One students for solutions to develop the application.

Application Development Timeline: when the researcher understood the needs of the users and various problems, then planned the application development timeline and created an application development plan.

The program operation format is as follows: 1) users use the application, 2) send data to the back-end using the API with voice calculation, 3) send data to store data in the database, 4) API selects data from the database, and 5) API sends data to the user.

Simulate database creation, simulating a database to be adapted for use in application development using PHP Myadmin, which is selected because it is a free database tool.

Creating an API and model, the researchers used FastAPI to create a data transmission line between the user and the back-end system because FastAPI is easy to use and can also fit well with our model, similar to Python.

To create an interface page and connect it to the back-end system, the researchers created an interface page for recording audio for analysis using the Vs Code program. There is a login page to display user information. User selection page to test the audio. Audio recording page to analyse audio for words. Audio recording page to analyze audio for words. The researchers created an interface page to display the results of the operation; when all audio was recorded, the program displayed the results. The results were displayed in a table format for each person to check again.

Suggestions

This Thai word analysis and accuracy test of Thai pronunciation of first-grade primary school students is a research on the synthesis of Thai words by categorising words according to their characteristics to make it easier to use words for testing and entering the application development process. This study also compared the similarity of sound waves by using models for training and testing. Therefore, it is necessary to study various theoretical data and techniques to select the appropriate technique for the development of a system with higher efficiency in sound signal similarity. Applications to help analyse Thai pronunciation should be developed to be diverse and appropriate for students' grade levels. In addition, research should be conducted to study the results of using applications to help analyse Thai pronunciation according to reading tests of words that are often mispronounced in Thai with students in other grades. Additionally, if this research technique is applied to other tasks, the accuracy and appropriateness of the analysis technique should be considered. Because there are many types of sound signal processing techniques, each technique has its own advantages and disadvantages. Therefore, information regarding each type of technique should be studied in detail to select the appropriate technique for future research.

References

Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel frequency cepstral coefficient and its applications: A review. *IEEE Access*.

Academic Division, Ministry of Education. (1937). *Siam grammar, prosody, Thai language textbook*. Ministry of Education.

Ahuja, L. (2025). Advanced speech recognition: Techniques, challenges, and applications. In A. Sharma & R. Rani (Eds.), *Artificial Intelligence and Speech Technology*.

Ali, S., Tanweer, S., Khalid, S., & Rao, N. (2021). Mel frequency cepstral coefficient: A review. *EAI Endorsed Transactions on Scalable Information Systems*, *27*(2).

Anchalinukul, S. (2003). *Thai word system*. Chulalongkorn University.

Anjali, I. P., & Sherseena, P. M. (2020). Speech recognition. *International Journal of Engineering Research & Technology (IJERT), 8*(4), 1-3.

Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission. (2011). *Report on the study of basic vocabulary of primary school students, Grades 1–3, Academic Year 2011* \

[Unpublished report]. Ministry of Education, Royal Thai Government.

Chanawong, P. (1986). *Thai grammar*. Nakhon Si Thammarat Teachers College.

Changkhwanyuen, W., et al. (2006). *Thai language standards, Volume 2: Words, word formation and word borrowing*. Ministry of Education.

Changkhwanyuen, P. (1982). *The art of listening and reading* (p. 4). Academic Press.

Chanon, W. (2007). *Effective reading*. Retrieved December 18, 2023.

Churchun, W. (1990). *The Result of using Songs in Drilling alphabet Pronunciation of Prathom Suksa Two Students*. Chulalongkorn University.

Kaewsanae, T. (1999). *How to read, read well* (3rd ed.). Thai Wattana Panich.

Kuhapinan, C. (1999). Reading and reading promotion. *Silpa Bannakarn*.

Laimoon, P., & Aroonmanakun, W. (2017). An analysis of commonly misspelled Thai words. *Humanities Journal*, *24*(2), 318-343.

Lamduan, S. (1983). *Thai grammar*. Odeon Store.

Lanchwathanakorn, C. (2016). Guidelines for assisting elementary children struggling with reading. *Journal of Education Studies, Chulalongkorn University, 44*(3), 287-301.

Malangpoo, P., Philuek, W., & Pomsamrit, N. (2022). Using speech recognition system for enhancing Chinese pronunciation. *International Journal of Mechanical Engineering, 7*(1), 3442-3451.

Moxon, S. (2021). Exploring the effects of automated pronunciation evaluation on L2 students in Thailand. *International Journal of Education, 9*(3), 41-57.

Muda, L., Begam, M., & Elamvazuthi, I. (2010). Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *Journal of Computing, 2*(3).

Nong Rangsit School. *(2021). Annual work plan of Nong Rangsit School*. Sukhothai: Nong Rangsit School Administration Group.

Office of the Basic Education Commission. (2025). *Testing management systems: Reading Test (RT) score report*. Retrieved March 20, 2025.

Passon, S., & Philuek, W. (2024). Effects of using applications to Thai phonetic recognition according to the Reading Test of Grade 1 students. *Journal of Computer and Creative Technology, 2*(3), 145-158.

Phraya-Upakitsanlapasarn. (1995). *Thai language principles: Spelling, verbology, syntax, and chants* (13th ed.). Bangkok: Thai Wattana Panich.

Philuek, W., & Puttasem, D. (2023). Teaching Thai language literacy: Proposal of using speech recognition technology techniques to detect read aloud in Thai tonal conjugation for primary education students. *Shanlax International Journal of Education, 11*(4), 77-84.

Phongphaiboon, S. (1988). *Thai grammar*. Bangkok: Thai Wattana Panich.

Phongprasert, P., & Makjui, A. (2021). The development of reading and spelling ability of Prathomsuksa 1 students taught by Hunter's teaching model and exercises (Doctoral dissertation, Silpakorn University, Thailand).

Premphan, F. (1999). *The science of Thai language use*. Kanchanaburi: Kanchanaburi Rajabhat Institute.

Ramakrishnan, S. (2012). Recognition of emotion from speech: A review. In *InTech*.

Royal Institute. (2011). *Royal Institute dictionary B.E. 2554* (2nd ed.). Bangkok: Royal Institute.

Sathira-angkool, N. (2003). *Thai pronunciation for effective communication*. Bangkok: Faculty of Education, Chulalongkorn University.

Sattayophas, S. (2002). *Thai language for communication and research*. Bangkok: 21st Century Company Limited.

Suchato, A., Pratanwanich, P. N., Chomphooyod, P., & Wiriyachaiporn, P. (2023). A study of the application of artificial intelligence to develop reading skills of primary school students.

Sukkeaw, J. (2022). *A development of reading ability on learning activities with brain based learning according to skills practice forms of Grade 1 students*. Naresuan University.

Sumrongpan, L., & Erawan, W. (2022). Development of a literacy promotion model in the Thai language learning strand for Grade 1. *Journal of Educational Measurement, Mahasarakham University, 28*(2), 222-237.

Teemuangsai, S., & Meesook, C. (2017). Thailand's classroom learning practices in secondary level: Are we ready for learning in the 21st century? *Global Educational Journal of Science and Technology, 8*(1), 1-12.

The Faculty of Thai Language, Thammasat University. (2000). *Thai language usage 1* (4th ed.). Bangkok: Thammasat University Press.

The Royal Gazette. (2024). Announcement of the Ministry of Education regarding the education policy of the Ministry of Education for the fiscal year 2025–2026. *Royal Gazette, 141*(Special Section 309), 11-13.

Thonglor, K. (1982). *Thai Grammar*. Bangkok: Bamrungsan.

Ubonyam, A. (1999). *Language for Communication*. Bangkok: Text and Journal Publications.

Vaidya, S. M., Hulawale, R. A., & Mohite, A. S. (2021). Speech recognition technology. *International Journal of Research Publication and Reviews, 2*(8), 771-774.

Waibanthao, W. (2020). *Learning Management by Using Tales to Develop Thai Reading Abilities of Grade 1 Students*. Rangsit University.

Warden, S. A., & Ye, Y. (2017). A comparative study of students' perceptions of expected ICT use and actual ICT use in classroom practices at Nawaminthrachinuthit Bodindecha School, Thailand. *Scholar: Human Sciences, 9*(1).

Author Details

Nicharee Pomsamrit, *The Faculty of Education, Nakhon Sawan Rajabhat University, Thailand,*
***Email ID**: nicharee@nsru.ac.th*

Siwadol Waraaeksiri, *The Faculty of Education, Nakhon Sawan Rajabhat University, Thailand,*
***Email ID**: Siwadol.w@nsru.ac.th*