ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 15, No. 1, pp. 294-306, January 2025
DOI: https://doi.org/10.17507/tpls.1501.32

The Effects of Text-to-Speech Dictation on Chinese Elementary Students' Listening and Spelling Skills at a Primary School in China

Suifang Yao
Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand

Kornwipa Poonpon*
Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand

***Abstract*—The objectives of the study were to investigate the effects of text-to-speech (TTS) dictation on Chinese elementary students' listening and spelling skills at a primary school and explore Chinese elementary students' perceptions of TTS dictation on listening and spelling skills. The present study adopted a mixed method. Quantitatively, the two-group pre-test and post-test design were employed. The 74 participants studied in Guangzhou Liwan Puiching Primary School (grade 4). The experimental group of 37 students where the TTS dictation was given as the intervention in the classroom, while the control group of 37 students where instruction was taught only in class. The quantitative data were collected through pre- and post-tests. A semi-structured interview method was used to collect qualitative data on student opinions. The results showed that the experimental group students performed better than the control group after the TTS dictation intervention. The average scores of different items in the experimental group were higher than those in the control group, indicating the impact of TTS dictation on Chinese primary school students' listening and spelling skills. The interview data also reveals that TTS dictation, as an intervention, sparked strong interest among most participants. Many of them reported that TTS dictation had a positive impact on their listening and spelling skills. The findings suggest that TTS dictation not only engaged the students but also contributed to noticeable effects in these key areas of listening and spelling skills.**

***Index Terms*—dictation, Text-to-speech dictation, listening skill, spelling skill, Text-to-speech**

I. Introduction

In the initial stage of English learning, learners mainly learn through listening and imitation (Amel & Sana, 2018). Therefore, many listening imitation opportunities and good external listening learning conditions are significant for learners. Teachers should provide and create such opportunities and conditions for learners to cultivate them to form correct pronunciation habits in the early stage (Gilakjani, 2011). However, there is generally non-standard voice input in the current English teaching. The input voice often has a personalised standardised voice that does not meet the requirements of English teaching. The phenomenon of English teachers' pronunciation not being standard enough or "dialect English" with accents has existed for a long time in some areas of our country. The nonstandard pronunciation and intonation of teachers' pronunciation will directly affect learners' listening and spelling. In the early stage of phonetic learning, once this non-standard pronunciation habit is formed, it is difficult to correct the mistakes of later students in dictation (Wen, 2022).

Dictation is a skill in which learners get some oral information in a short time and then write what they hear (Oller, 1971). English dictation is not only an important exam method but also an active and creative learning process (Deng & Zhang, 2004). It can accurately reflect students' overall language level and investigate the language ability and knowledge of students' listening, speaking, reading, and writing. With the advance of technology in vocabulary learning and teaching, dictation has evolved to be technology-based. Text-to-speech (TTS) was introduced as an effective technology-based tool to convert text input into speech output or read digital text aloud. You can choose an English accent (British, American, or others), voice (male or female), and speed (very slow to very fast). It can be used to improve listening and spelling skills in English (Chiang, 2019). TTS was applied to be used with dictation by inputting the dictation text into TTS and selecting it according to the required English accent, voice, and speed. TTS dictation helps teachers with low confidence in teaching contexts where teachers' pronunciation or accents are not clear during face-to-face dictation, resulting in students' inability to write correct words. Through TTS dictation, students can improve their listening and spelling skills (Huang & Liao, 2015). In particular, TTS dictation should be helpful to students who have just started learning English or are in primary school in developing their English listening and spelling skills. Zhu (2016) emphasized that applying TTS to English listening and spelling skills has a certain potential

* Corresponding Author. Email: korpul@kku.ac.th

for development. Since TTS dictation is widely used to develop English listening and spelling skills worldwide, some studies use TTS dictation to improve EFL learners' listening and spelling skills. Zhu (2023) stated in the study that the TTS dictation intervention positively enhances students' listening and spelling skills. In a study conducted by Kataoka (2007), high school students in Japan spelt words and phrases better by text-to-speech dictation, and they learned and memorised more words than without using TTS dictation. For elementary school students, especially in the context of English learning, TTS dictation provides crucial support in developing their listening and spelling skills.

II. Literature Review

A. Dictation in Teaching Listening and Spelling

Dictation has been widely used in teaching listening and spelling. Whitaker (1976) believes that dictation is a useful teaching method. Dictation highlights the importance of phonetics, helping students recognise and differentiate between similar sounds, which is essential for accuracy in both listening and spelling. As students listen and transcribe what they hear, they practice spelling, reinforcing the correct spelling of words by linking the sounds they hear to their written forms (Chu, 2019). Valette (1964) believed that teachers focus on diverse components of language, including sound and sentence structure, and urge students to edit their spelling and dictation, which might stimulate students' awareness of spelling. Teachers can offer immediate feedback during or after dictation exercises, enabling students to correct their mistakes and learn from them, thereby reinforcing both their listening and spelling skills.

For non-native teachers to use dictation in their classes, one of the major concerns could be clear pronunciation, as it is important for students to correctly understand what teachers say in the process of dictation. When students practice listening and spelling, they need to listen to clear and correct input. If the students cannot catch the words and sentences clearly, they are unlikely to spell the words correctly. This means that students' listening is somehow linked to teachers' pronunciation (Gilbert, 1995). Using clear and correct pronunciation in dictation is, therefore, important because it will help avoid the risk of rigidity and stability of pronunciation habits from the beginning (Bhatia & Ritchie, 2008). But how can non-native teachers who have learned English as a foreign language do this? With the help of technology these days, one method to do this is to use TTS to help in pronunciation and teaching listening and spelling (Kilickaya, 2006; Levis, 2007; Seferoğlu, 2005).

B. Text-to-Speech (TTS)

Text-to-speech (TTS) technology is a system that converts written text into spoken language, allowing users to listen to the synthesized speech output. This technology processes text input by analyzing the linguistic features and converting them into audible speech using algorithms and pre-recorded voice data.

The intelligibility of TTS speech output plays a crucial role in listening comprehension (Sha, 2010). Clear and accurate pronunciation enables words to be heard distinctly, aiding in correct spelling. Studies have shown the benefits of TTS technology in language learning contexts. For instance, Kataoka's study in 2007 revealed that high school students in Japan had better word and phrase memorization when exposed to text-to-speech voices, leading to enhanced English vocabulary acquisition compared to traditional methods.

While research on TTS technology is still limited, preliminary findings underscore its significance, particularly in improving listening and pronunciation skills in language classrooms. Zhang (2020), for instance, has successfully integrated TTS technology into English teaching practices, yielding positive educational outcomes. Li (2003) utilized TTS to simulate real-life English scenarios with accurate pronunciation, enhancing language learning experiences. Van Laere and van Braak (2017) emphasized TTS's role in supporting listening activities by addressing learners' challenges with non-standard pronunciations, thereby boosting listening proficiency. Sha (2010) developed TTS-based audio materials for listening comprehension, showcasing the technology's effectiveness in enhancing students' listening skills through natural and standard pronunciation. Furthermore, Kong and Lin (2001) highlighted TTS's ability to adjust speech speed and voice characteristics, catering to learners' diverse needs and levels of proficiency. By providing standard pronunciation and diverse listening experiences, TTS helps learners improve their listening skills and spelling accuracy, ultimately enhancing overall language proficiency.

C. Text-to-Speech (TTS) Dictation in Chinese Elementary Students' Listening and Spelling Skills

Text-to-speech (TTS) dictation is defined as the teacher selecting the appropriate text content to submit to the TTS system, which analyses and processes it to build a speech. After the students input the speech, they write down the content. According to Nation's (1991) dictation framework and the TTS framework by Dutoit (1997) and Sasirekha and Chandra (2012), the TTS dictation framework is illustrated in Figure 1.

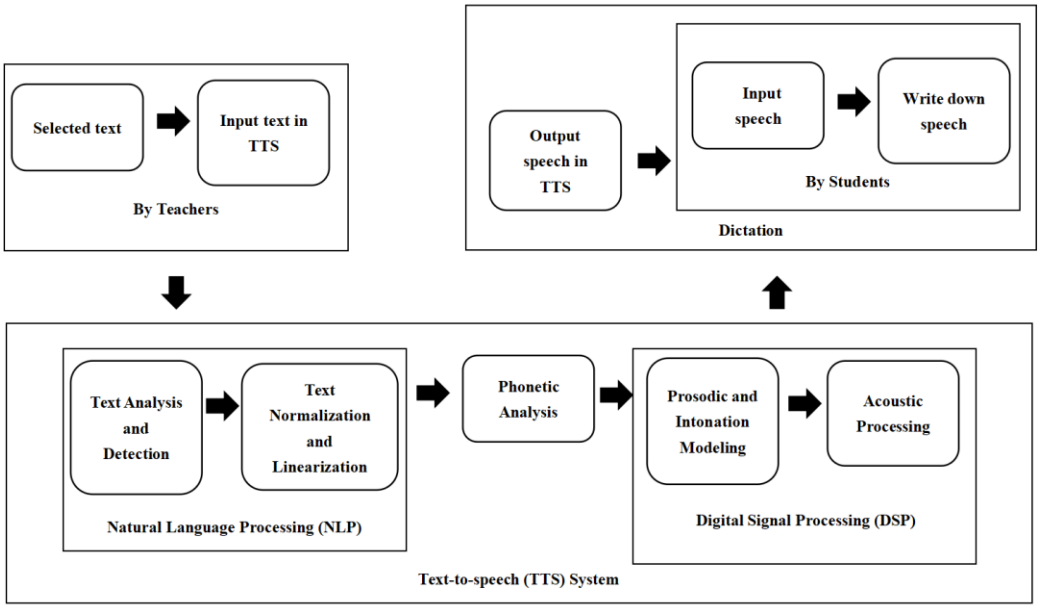

Figure 1. The Framework of TTS Dictation

According to Figure 1, the TTS system consists of three parts: natural language processing (NLP), phonetic analysis, and digital signal processing (DSP). Natural language processing (NLP) could create phonetic symbols of the text to be read and the required intonation and rhythm. Natural language processing (NLP) consists of two parts: 1) text analysis and detection, and 2) text normalization and linearization. Digital signal processing (DSP) turns the received symbol information into spoken speech. Digital signal processing (DSP) consists of two parts: 1) prosodic and intonation modelling, and 2) acoustic processing. After the voice output in TTS dictation, students input the voice (listening input) and write down the heard content (writing output) to form a dictation process.

Xu and Dai (2010) mentioned that teachers can now create various documents (e.g., PowerPoint, Word, PDFs, or Texts) to provide different TTS inputs, and the process is easy to operate. Some TTS systems can even download MP3 or access it offline. After the teacher uploads the text, TTS generates a voice after analysis and processing. Teachers can choose accent (British/American English), voice (male/female), and speed (very slow to very fast) before playing voice to students. If teachers must dictate many times, they can even click the text several times. After listening to the voice, students write down the content read by the voice and complete the dictation process.

In China, English is used as a foreign language for students. Affected by the economic development level of different regions in China, many English teachers have not had the opportunity to go abroad to receive education in the native English language, and even have almost no exposure to audio/video resources of British/American pronunciation. This has also led to certain differences in pronunciation between Chinese English teachers themselves and English teachers whose native language is English. Their pronunciation will immediately affect their English teaching and even affect students' dictation. When a teacher mentions a word, but its pronunciation is not standard, it may lead to students writing the wrong word, affecting their English learning. Helping teachers use clear and correct pronunciation in dictation is a daunting challenge for most English teachers in China. Lack of government funding or technical equipment, lack of technical skills among teachers, and lack of English/American language input are just a small part of the various problems faced by English teachers. TTS dictation has the potential to help teachers solve this problem. TTS dictation is an aspect of TTS in teaching, but it can solve the problems caused by teachers' non-standard pronunciation and uneven dictation speed. TTS dictation can be used not only for face-to-face courses but also for online courses, greatly improving students' listening and spelling skills.

TTS dictation allows students to practice listening and spelling skills, including sentence and vocabulary drills. However, when teachers do not have enough time to enhance pronunciation and schools cannot provide corresponding aid, TTS dictation is surely an excellent concept for training students' listening and spelling. TTS can be coupled with dictation in the English process (González, 2007). Only with appropriate pronunciation can pupils write down the right words and improve the effectiveness of vocabulary learning (Zhu, 2016).

Consequently, the study was guided by the following two research questions.

1. What are the effects of TTS dictation on Chinese elementary students' listening and spelling skills in Guangzhou Liwan Puiching primary school?

2. What are Chinese elementary students' perceptions of TTS dictation on listening and spelling skills?

III. METHODOLOGY

A. Research Design

This study employed a mixed-method research design, integrating both qualitative and quantitative approaches to provide a comprehensive understanding of the effects of TTS dictation on students' listening and spelling skills. The combination of these methods allowed for a more nuanced analysis of the data, offering both depth and breadth in the research findings. Figure 2 shows the mixed-method design of this study. First, two instruments were designed: an instructional instrument and a research instrument. Second, the quantitative data from the pre-test and post-test were compared with qualitative insights from semi-structured interviews to identify any discrepancies. Finally, the results from both the quantitative and qualitative phases were interpreted together to understand not only the effects of the TTS dictation intervention but also the students' perceptions of its impact on their listening and spelling skills.

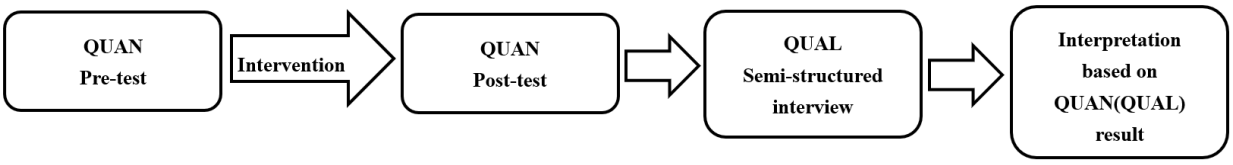

Figure 2. The Mixed-Method Design of This Study (Creswell & Clark, 2007)

B. Participants

The participants of this study were Grade 4 students enrolled in an English course during the 2023 academic year. The 74 participants were selected using a convenience sampling method because the researcher was teaching at this school and had to conduct the study with intact classes. The students were informed about the research project and gave their consent prior to the study. Ethical approval for this study was also obtained from the Center for Ethics in Human Research, Khon Kaen University (APPROVAL NUMBER: HE663132).

C. Instruments

The study used two instruments: the instructional and the research instruments. The instructional instrument was lesson plans. Test papers (pre- and post-test) and semi-structured interviews were used as research instruments. The test paper was to collect quantitative data, and the semi-structured interview was used to collect qualitative data.

(a). Instructional Instruments

Lesson Plans

The lesson plans aimed to provide teaching content and methods that ensured students could actively participate and learn through TTS dictation, leading to successful outcomes in the post-test. These plans were designed for a 6-week instruction period. The research was based on the TTS dictation instructional framework recommended by the Guangzhou Municipal Teaching and Research Office. A teaching plan covering 4 units was developed accordingly.

The Ministry of Education of Guangzhou designated the contents and must be studied by students. The students learned the following contents: 1) the words of Unit 1- Unit 4; 2) the text (dialogue) of Unit 1- Unit 4. It consisted of Unit 1: What's In Your Room?; Unit 2: They Are Near the Window; Unit 3: Welcome to My House.; Unit 4: I Live In a Big House. The lesson plans for TTS dictation instruction specify that all settings were configured during class.

All lesson plans were based on the TTS dictation framework, which includes four stages: previewing vocabulary, reviewing vocabulary, previewing text (dialogue), and reviewing text (dialogue). During the vocabulary preview stage, students engaged in activities to familiarize themselves with the vocabulary for each unit. They learned the meanings and usage of the words, practised pronouncing them in class, and focused on phonetics, especially vowels, and consonants. After learning, students completed spelling exercises and checked their answers. At the review stage, students went over the words they had previously learned, focusing on pronunciation and their Chinese meanings. After the review, they completed a TTS dictation and submitted it to the teacher for feedback. The teacher corrected their work and provided feedback, and the students promptly corrected their mistakes, reinforcing the correct words in class.

At the text (dialogue) preview stage, students learned the text for each unit and discussed the vocabulary they had learned. After this, they completed another TTS dictation, submitted it for feedback, and followed the same process of correction and reinforcement. During the text review stage, students reviewed and memorized the text and vocabulary for each unit. They then did another TTS dictation, received feedback, corrected their mistakes, and reinforced the correct words in class.

These lesson plans were evaluated and approved by experts from the Liwan District Education Research Institute of Guangzhou and were adjusted according to their suggestions before being implemented in practical research.

(b). Research Instruments

1. Test Papers

In the current study, two tests, a pre-test, and a post-test, were developed from the same test blueprint. They were used to measure students' listening and spelling before and after the TTS dictation intervention. The two tests consisted of five parts: 1) Listen and Choose, 2) Listen and Order, 3) Listen and Spell, 4) Listen and Tick or Cross, and 5) Listen and Fill the text. This test was adapted from the 2020-unit test issued by Liwan District Education Research Institute of

Guangzhou. The test required the students to listen to each item three times and write the answer after listening. The test paper contained a total of five parts with a 100 total score: ten items in part 1 with 20 points, 8 items in part 2 with 16 points, 4 items in part 3 with 8 points, 8 items in part 4 with 16 points, and 5 items in part 5 with 40 points. The test lasted 40 minutes.

2. Semi-Structured Interviews

Qualitative data was collected by semi-structured interviews to explore Chinese elementary students' perceptions of TTS dictation on listening and spelling skills. At the end of the experiment, a semi-structured interview was conducted with 10 students (5 with the highest test scores and 5 with the lowest test scores) in the experimental group to investigate the perception of the TTS dictation in the experiment. Due to the under-18 years old of the participants, an English teacher who taught other grades in the school conducted semi-structured interviews and audio recordings of 10 participants (the approval of the student's parents and students was secured in advance). In this process, the student's parents were alerted by phone and requested, that one party be present to observe. Questions were adapted from Cheng (2018). The semi-structured interview guided consists of opening, interviewing, and closing. Responses are coded and analyzed using content analysis (Creswell, 1998). These questions were translated into Chinese. All interviews were conducted in Chinese to avoid participants' misunderstanding and facilitate students' perception expression. Due to the similar answers from the data collection by audio recording, each question from the semi-structured interview was divided and investigated.

D. Data Collection

The data collection was conducted during the first semester of the academic year 2023 (September - October 2023). The time and method of data collection procedures in terms are depicted in Figure 3.

According to Figure 3, the participants were orientated about the objectives of the study and asked for their willingness to participate in the study. Next, the pre-test was given to the participants by the researcher. Then, the TTS dictation intervention (units 1-4) was used, respectively. Next, the instruction was taught following the lesson plans by stages: preview vocabulary, review vocabulary, preview text, and review text. After that, the researcher administered the post-test to the participants. Based on the ranking of scores, the five highest-scoring participants and five lowest-scoring participants were selected in a semi-structured interview. Finally, the researcher conducted a semi-structured interview to better understand the perceptions of TTS dictation on listening and spelling skills by the participants.

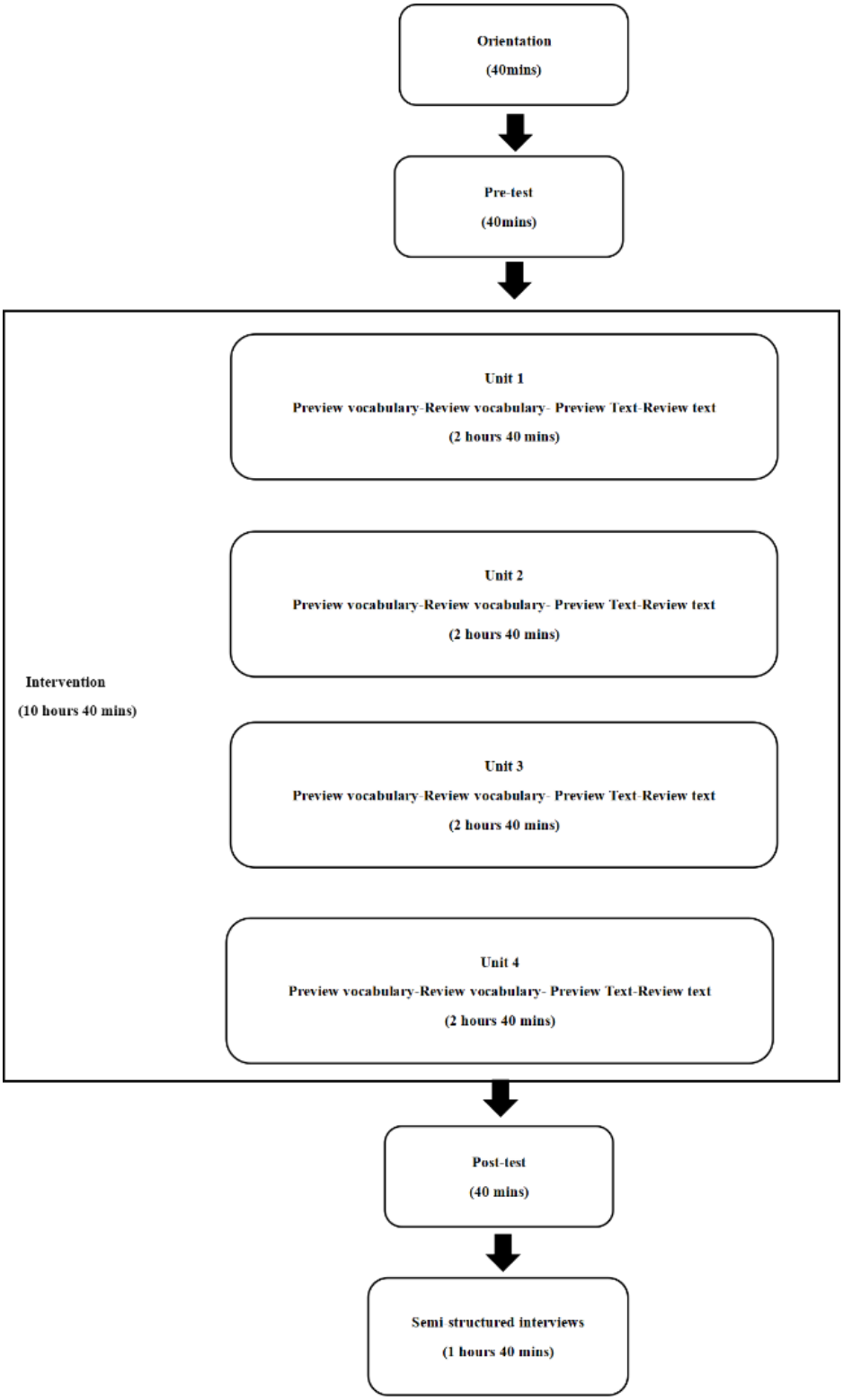

Figure 3. Data Collection Procedures

E. Data Analysis

The scores of both the experimental and control groups on each part of the test paper (pre- and post-test) were analyzed using SPSS. To assess the differences within the experimental group, a paired samples t-test was conducted on the pre-test and post-test scores. To compare the scores between the control and experimental groups, an independent samples t-test was used. The results provided a clear understanding of the students' listening and spelling skills, as indicated by their test scores.

Through the above tests, the students' listening and spelling skills were fully understood. The test results were to tell students the effects of listening and spelling through the grades. After testing, Table 1 shows the grading standard. Students with scores of 85-100 can achieve A level, scores of 75-84.5 can achieve B level, scores of 60-74.5 for C level, and scores below 60 are considered D level. The grading standard is issued by the Guangzhou Liwan District Education Research Institute, and the Grade 4 teachers correct all the papers and inform the students of the grades (A, B, C, and D respectively) (China compulsory education primary school English curriculum standards, 2011).

In addition to the quantitative analysis, qualitative data were gathered through semi-structured interviews. The researcher observed and recorded the participants' perceptions of the TTS dictation's effects on their listening and spelling skills. These responses were coded by two experienced English teachers using the content analysis framework developed by Kondracki and Wellman (2002). The coders identified both positive and negative perceptions held by the participants regarding the TTS dictation intervention. The inter-rater agreement between the coders was 90%, ensuring reliability in the analysis of students' perceptions.

TABLE 1
THE GRADE STANDARD FOR PRIMARY SCHOOL STUDENTS IN LIWAN DISTRICT, GUANGZHOU (CHINA COMPULSORY EDUCATION PRIMARY SCHOOL ENGLISH CURRICULUM STANDARDS, 2011)

Grade	Score
A	85-100
B	75-84.5
C	60-74.5
D	Below 60

IV. RESULTS AND FINDINGS

A. *The Effects of TTS Dictation on Chinese Elementary Students' Listening and Spelling Skills*

To observe if there was a significant difference in the pre-test and post-test mean scores of participants in the experimental group (EG), the paired sample t-test was used. Table 2 indicated a significant difference between the participants' test scores in the EG, t (36) =12.37, p<.05.

Table 3 displayed the pre-test and post-test scores in the control group (CG) and experimental group (EG), showing the minimum score, maximum score, mean, and standard deviation. The pre-test scores showed little difference between EG and CG in the minimum, maximum, mean, and standard deviation data. The post-test mean scores in both groups were higher than the pre-test mean scores. The post-test of the SD in both groups was lower than the SD of the pre-test. As shown, the post-test scores of the EG (M=94, SD=8.60) were significantly higher than the pre-test scores (M=36.35, SD=24.63), the mean score of the EG was higher than the CG in the post-test, which revealed that the listening and spelling skills of students in the EG were improved during the TTS dictation intervention.

TABLE 2
THE PAIRED SAMPLES T-TEST OF THE PRE-AND-POST-TEST MEAN SCORES

Paired Differences								
Test papers		M	SD	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Pre-test	Post-test	57.649	28.342	48.199	67.098	12.372*	36	<.001

*P<0.05

TABLE 3
THE PRE-AND-POST-TEST MINIMUM, MAXIMUM, MEAN SCORES AND SD OF THE PARTICIPANTS IN THE CONTROL AND EXPERIMENTAL GROUP

Test papers	Groups	Minimum	Maximum	Mean	Std. Deviation
Pre-test	CG	4	91	36.92	26.857
	EG	6	94	36.35	24.639
Pro-test	CG	56	100	87.19	12.023
	EG	70	100	94.00	8.602
Different score	CG	52	9	50.27	28.235
	EG	64	0	57.65	28.342

Furthermore, an independent samples t-test was conducted to determine if there was a significant difference in the mean scores between the experimental group (EG) and the control group (CG). Cohen's d was calculated to assess the effect size of this difference. The results revealed a significant difference between the post-test mean scores of the participants in the control group (M=87.19, SD=12.02) and those in the experimental group (M=94, SD=8.60), t (98.78)=-16.68, p<.001. This indicates that the TTS dictation intervention had a significant positive impact on the listening and spelling skills of the participants in the experimental group.

As shown in Table 5, a distinctly higher number of participants in the experimental group achieved an A-level in the post-test compared to the control group. This further supported the conclusion that the listening and spelling skills of the participants in the experimental group were improved during the TTS dictation intervention.

TABLE 4
THE INDEPENDENT SAMPLES T-TEST OF THE PRE-AND-POST-TEST SCORES AND THE DIFFERENT MEAN SCORES

Independent Samples Test					
Test papers	M	SD	t	df	Sig. (2-tailed)
Pre-test	-.57	2.24	-.095*	72	.925
Post-test	6.81	3.42	2.802*	72	.007
Different score	6.24	-.11	-16.68*	98.78	.001

*P<0.05

TABLE 5
THE POST-TEST GRADES OF THE EXPERIMENT GROUP AND CONTROL GROUP

Grade	EG		CG	
	N	Percentage	N	Percentage
D= 0-59.5	0	0%	2	5.4%
C=60-74.5	3	8.1%	4	10.8%
B=75-84.5	2	5.4%	6	16.2%
A=85-100	32	86.5%	25	67.6%
Total	37	100.%	37	100.%

B. The Chinese Elementary Students' Perceptions of TTS Dictation on Listening and Spelling Skills

The top 5 and bottom 5 participants in the experimental group took part in a semi-structured interview regarding their perceptions of the TTS dictation. The participants' perceptions were collected and categorized according to the four main ideas derived from the interview questions.

First, the participants believed that it was difficult to memorize English words and then used different memorization methods. Most of the participants in EG considered it difficult to memorize the words.

For example:

"*I would forget them by the next day*." (S1)
"*English isn't my first language*." (S3)
"*It was difficult to remember words before*." (S4)
"*It's not as easy for me to remember things like it is with my native language.*" (S5)
"*I hardly ever come across English in my daily life, so outside of class, it's tough to remember words.*" (S7)
"*Being unfamiliar with the pronunciation of English words makes it difficult to memorize them.*" (S8)
"*The difference between the English and Chinese writing systems makes it difficult for me to adapt to spelling or recognizing words, which in turn makes it challenging to remember them.*" (S10)

To deal with this challenge, the students used different methods to remember words. They used rote memorisation, class reading memorisation and duplicate hand-writing memorisation.

For example:

"*I used to rely on rote memorization*." (S6)
"*I wrote the words over and over to try and remember them.*" (S9)
"*I used to try to remember how the teacher pronounced words in class and then review them at home*." (S2)

Second, the participants in the interview recognized the necessity and advantages of dictation in English learning. Most of the participants believed that dictation could be an effective method for memorizing new words and improving their listening and spelling skills.

For example:

"*I thought dictation is necessary and useful.*" (S2)
"*Through dictation, I could focus on what I hear and then write it down, which helps me remember the words.*" (S5)
"*I wrote down what I heard, which helped me practice spelling.*" (S7)
"*It deepened my understanding of the correct spelling of words.*". (S9)
"*While listening, I focused more on the correct pronunciation of the words, which helped me spell them accurately.*" (S10)

Some participants believed that dictation provided authentic language contexts, which helped them understand how words were used in the sentences or passages. They also could receive immediate feedback from teachers, making it easier to remember them and use them correctly in the future.

For example:

"*I learnt the usage of words within sentences or texts.*" (S3)
"*I knew how to use words in real-life situations.*" (S4)
"*My teacher could provide me with immediate feedback after the dictation, allowing me to quickly identify and correct my mistakes.*" (S1).
"*I knew which words I spelt right in immediate feedback and which I could not*." (S6)

On the other hand, a small number of students had negative perceptions of dictation.

For example:

"*Dictation was not necessary in class. I didn't like dictation.*" (S8)

The third aspect of the study focused on the participants' favour for TTS dictation. The results revealed that 9 EG participants favoured TTS dictation in English class, and only 1 participant did not prefer dictation and did not favour TTS dictation.

For example:

"*I liked TTS dictation. But if I don't pay close attention, I end up spelling the words incorrectly*." (S6)

"*I didn't like dictation and spelling.*" (S8)

Those who liked TTS expressed a preference for the flexibility in adjusting the speed of dictation and appreciated the standard pronunciation provided by the TTS dictation. They found that being able to control the speed allowed them to follow along more comfortably, and the clear, consistent pronunciation helped them better understand and retain the words they were learning.

For example:

"*The speed of TTS dictation is very suitable for us. During TTS dictation, the teacher would adjust speed according to my dictation, but I didn't get all the right answers in the exam because I wrote slowly and couldn't keep up.*" (S9)

"*The pronunciation in TTS dictation was very clear and similar to the English stories or recording I listened to at home, so it was easy for me to write the words down.*" (S1 and 4)

"*The pronunciation was clear and loud.*" (S2)

"*The speed of TTS dictation is just right for me, and the teacher can slow it down based on my needs.*" (S7)

Compared to the teacher's pronunciation in class, some participants expressed favour for the standard pronunciation provided by TTS dictation.

For example:

"*The pronunciation in TTS dictation was clearer and more accurate than the teacher's, making it easier for me to spell words correctly.*" (S10)

"*The pronunciation of TTS dictation was very standard, which was clearer than the teacher's pronunciation, especially the final sounds, so I could write it easily.*" (S3)

"*I preferred using TTS dictation than the teacher's pronunciation because it is more accurate.*" (S5)

Finally, after 6 weeks of TTS instruction, the students reported significant improvement in their listening and spelling skills as a result of the TTS dictation. For example:

"*The pronunciation in TTS dictation was very clear, which allowed us to accurately write the words.*" (S2)

"*Sometimes the teacher overlooked some of the final sounds, but we were able to hear them clearly through TTS dictation.*" (S3)

Some students mentioned that the standard pronunciation provided by TTS dictation helped them achieve better scores on exams.

For example:

"*The pronunciation of the words in the TTS dictation was almost identical to the pronunciation in the test, and I was able to write them down.*" (S1)

"*It improved my performance in the test. By using TTS dictation, I could find corresponding answers to many listening questions, which increased my accuracy.*" (S4)

"*Using TTS dictation made me more familiar with the pronunciation of certain words. When I listened to the recording, I could immediately write down the word.*" (S5)

"*Compared to the teacher's dictation before, TTS dictation significantly improved my listening and spelling scores.*" (S6 and 7)

"*Although I didn't achieve very high grades, I felt that my listening and spelling skills improved significantly with TTS dictation.*" (S9)

"*My listening and spelling skills improved significantly in the exam.*" (S10)

Nevertheless, few students expressed negative perceptions of TTS dictation.

For example:

"*TTS dictation was just a simple playback of sound. I don't know how it helped me.*" (S8)

The semi-structured interviews with the top five and bottom five participants from the experimental group provided valuable insights into their perceptions of the TTS dictation intervention. Most participants acknowledged the

challenges they faced in memorizing English words and employed various memorization techniques, such as rote memorization, class reading, and repetitive writing, to overcome these difficulties. Despite these challenges, a majority of the students recognized the effectiveness of dictation in improving their listening and spelling skills. They appreciated the focus on pronunciation and the ability to write words accurately, as well as the immediate feedback from teachers, which helped reinforce their learning.

Many participants expressed a clear preference for TTS dictation, citing the benefits of its standard pronunciation and adjustable speed, which made the learning process more manageable and effective. The clarity and consistency of the TTS pronunciation were particularly valued, with some students noting that it was easier to understand and follow than the teacher's pronunciation. As a result, students reported significant improvements in their listening and spelling abilities over the 6-week period, with several attributing their better exam performance to the use of TTS dictation.

However, it is important to note that a small number of students had negative perceptions of TTS dictation, finding it unhelpful or viewing it as just a simple playback of sound without a clear connection to their learning improvement. Despite these mixed reactions, the overall feedback suggests that TTS dictation played a positive role in enhancing students' English language skills, particularly in listening and spelling.

Many of the participants expressed that the regular practice of listening to the text and spelling out the words helped them better understand pronunciation patterns and word structures. This consistent exposure allowed them to internalize the correct spelling of words and improved their ability to recognize and reproduce these words accurately in different contexts. The students also noted that the interactive nature of TTS dictation made the learning process more engaging, which contributed to their increased motivation to focus on the material and retain the information more effectively. Overall, the TTS dictation not only reinforced their spelling and listening abilities but also boosted their confidence in using these skills in both written and spoken English.

V. Discussion

The findings of this study demonstrated that during the TTS intervention, the participants' listening and spelling skills were significantly improved. Most participants expressed positive perceptions towards the TTS dictation. The major findings of two research questions are discussed in this section.

The results of this study showed that the mean score of participants in the experimental group (EG) was significantly higher than that of the control group (CG). The improvement in participants' listening and spelling skills in the EG can likely be attributed to the TTS dictation intervention. Most participants indicated that they perceived the effect of TTS dictation on their spelling and listening skills to be positive, with noticeable improvements in both areas.

A. A Chance to Leverage the Integration of Information Technology and Traditional Learning Tools in Chinese Primary School Classrooms

This study highlighted the potential for integrating TTS technology with traditional learning tools in Chinese primary school classrooms. While previous research had emphasized the importance of a teacher's pronunciation in dictation exercises (Chiang, 2019), this study suggested that TTS technology, which used standardized British/American pronunciations, could offer an effective alternative to non-native English teachers' pronunciations. Zhu (2016) also noted the developmental prospects of applying TTS in enhancing English listening and spelling skills. However, there was a noticeable gap in research on the use of TTS in dictation exercises, particularly in the context of primary education in China (Huang & Liao, 2015). This study contributed to filling this gap by exploring the combination of TTS and dictation, an area that had been under-researched.

The findings of this study were consistent with several previous studies. For example, Pellegrini et al. (2012) found that students were more receptive to TTS-generated utterances than to human-transcribed ones in dictation exercises for European Portuguese learners. Similarly, Chiang (2019) demonstrated that TTS dictation positively impacted vocabulary performance among Taiwanese college students, with significant differences observed between TTS dictation and traditional teacher-led dictation (TLD).

In contrast, Kazazoğlu (2013) found that participants made fewer errors in TLD than in recorded dictation, suggesting that the "speed" of delivery might be a key factor affecting short-term memory retention. Despite this, the flexibility of TTS dictation—where the speed and type of voice could be adjusted—might have provided a more adaptable tool than TLD, offering distinct advantages in language learning.

The initial hypothesis suggested that the integration of TTS technology with traditional dictation methods would result in improved listening and spelling skills among students. The results of this study supported this hypothesis, demonstrating that TTS dictation could serve as an effective intervention in enhancing these skills. The findings indicated that combining TTS as an IT tool with traditional dictation methods created a new, effective learning intervention for Chinese elementary students. This combination leveraged the strengths of both approaches, offering significant benefits for vocabulary acquisition and overall English learning.

In summary, this study's findings confirmed that TTS dictation could be a valuable tool in English language education, particularly when combined with traditional learning methods. The positive impact on students' listening and spelling skills, as well as the favourable perceptions of TTS dictation among participants, suggested that this approach could be effectively integrated into the curriculum for primary school students in China.

B. A Chance to Improve Chinese Elementary Students' Self-Confidence in English Learning

For Chinese primary school students, English poses a significant challenge as it is learned as a second language. The findings of this study align with previous research, such as Huang and Liao (2015), who found that TTS dictation improved students' intrinsic motivation to learn English. Mispronunciation by teachers can lead to incorrect word usage, which in turn affects students' confidence in learning English. Zhu (2016) also highlighted that TTS dictation helps students overcome these pronunciation challenges, thereby enhancing their confidence in English learning by improving vocabulary acquisition and accuracy.

The results of this study support these earlier findings, indicating that TTS dictation not only enhances listening and spelling skills but also boosts students' self-confidence in learning English. This study further corroborates Chu (2019), who emphasized that dictation as an intervention provides learners with valuable input and feedback, essential for developing listening and spelling skills.

The initial hypothesis posited that TTS dictation would positively impact students' listening and spelling skills, leading to higher academic performance and improved perceptions of English learning. The results confirmed this hypothesis, as the experimental group showed significant improvement in both areas compared to the control group. Moreover, the participants' feedback reinforced the hypothesis that TTS dictation could be an effective tool in English learning, particularly in fostering word acquisition and building confidence.

Surprisingly, the qualitative data revealed an unexpected finding: most participants not only found dictation necessary for English learning but also preferred TTS dictation over traditional teacher pronunciation. They believed that TTS dictation was more reliable and helped them spell and understand words more accurately.

This study's findings support the integration of TTS dictation as an innovative approach in Chinese elementary classrooms. The combination of TTS technology and traditional dictation not only improves students' listening and spelling skills but also enhances their confidence in English learning. Future research should continue exploring this combination to further validate its effectiveness and expand its application in different educational contexts.

Some limitations of this study should be noted. The study used a small sample size, which means the results may not be generalizable to similar populations. Therefore, further research with larger and more diverse samples is needed to confirm these findings and allow for broader generalization. Future studies could explore the long-term effects of TTS dictation on different aspects of language learning, such as writing and speaking skills, and investigate its effectiveness across various age groups and proficiency levels. Additionally, research could examine the impact of TTS dictation when combined with other digital learning tools to enhance overall language acquisition. Further research was recommended to explore this combination's long-term benefits and potential applications in different educational contexts.

VI. Conclusion

This study explored the effects of TTS dictation on the listening and spelling skills of Chinese elementary students, finding that this technology-enhanced method significantly improved these skills. The results demonstrated that participants in the experimental group, who used TTS dictation, outperformed those in the control group in both listening and spelling tasks. The majority of participants also reported positive perceptions of TTS dictation, noting its effectiveness in helping them acquire new vocabulary and improve their pronunciation.

The findings align with previous research, highlighting the potential of integrating information technology, such as TTS, with traditional learning tools like dictation. This combination offers a promising new approach to vocabulary learning, particularly in contexts where English is taught as a second language. The findings of this study suggest that integrating TTS dictation into the English learning curriculum could enhance the effectiveness of language acquisition among Chinese elementary students. Not only does TTS dictation provide accurate pronunciation, but it also engages students in a way that traditional methods may not. This has the potential to boost students' confidence in their English skills, ultimately leading to better learning outcomes.

However, the study's small sample size limits the generalizability of the results. Future research should involve larger and more diverse populations to validate these findings and explore the long-term effects of TTS dictation on various language skills. Additionally, further studies could investigate the effectiveness of combining TTS dictation with other digital learning tools to create a more comprehensive language learning experience.

In conclusion, TTS dictation shows great potential as an innovative instructional tool in English language education, particularly for enhancing listening and spelling skills. As technology continues to evolve, its integration into language learning should be further explored and refined to maximize its benefits for Chinese students.

Acknowledgements

I would like to express my deepest gratitude to all the participants who contributed their time and effort to this study. Their cooperation was invaluable to the success of this research. I am also sincerely grateful to my university for providing the resources and support necessary to conduct this study. The guidance and encouragement from my professors and colleagues were instrumental in completing this work. A special thanks goes to Khon Kaen University (KKU) for its academic and administrative support throughout my research journey. The opportunity to collaborate with

KKU has been a significant factor in the development and completion of this thesis. Thank you to everyone who made this research possible.

REFERENCES

[1] Amel, O. U. A. R. E. T., & Sana, L. A. T. R. E. C. H. E. (2018). *The Effect of Listening and Imitation on Learners' Consonants Pronunciation* (Doctoral dissertation, Abdelhafid Boussouf University Centre-Mila).

[2] Bhatia, T. K., & Ritchie, W. C. (Eds.). (2008). *The handbook of bilingualism*. John Wiley & Sons.

[3] Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five approaches*. Retrieved Nov. 5th ,2024, from https://books.google.com/books/about/Qualitative_Inquiry_and_Research_Design.html?id=Ykruxor10cYC

[4] Creswell, J.W. & Plano Clark, V.L. (2007). *Designing and conducting mixed methods research*. Sage.

[5] Cheng, H.Y. (2018). *An empirical study on the effectiveness of the dictation strategy based on the Law of Forgetting in English teaching in Senior One* (Master's thesis, Yili Normal University). China National Knowledge Infrastructure. Retrieved Nov. 5th ,2024, from https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201802&filename=1018080416.nh

[6] Chiang, H. H. (2019). A comparison between teacher-led and online text-to-speech dictation for students' vocabulary performance. *English Language Teaching*, *12*(3), 77-93.

[7] Chu, T. (2019). The effect of task repetition with dictation intervention on the development of Chinese university minority EFL students' spoken English: A case study of Honghe University in Yunnan Province. *Asian Journal of Literature, Culture and Society*, *9*(1), 79-79.

[8] Dutoit, T. (1997). *An introduction to text-to-speech synthesis* (Vol. 3). Springer Science & Business Media.

[9] Deng. Y., & Zhang, W. Z. (2004). Analysis of the Use of Listening Strategies in English Dictation. *Journal of Xi 'an International Studies University*, *2004*, *12*(2), 61-65.

[10] Gilbert, J. (1995). Pronunciation practices as an aid to listening comprehension. In D. J. Mendelson & J. Rubin (Eds.), *A guide for the teaching of second language learning* (pp. 97–111). Dominic Press.

[11] González, D. (2007). Text-to-speech applications used in EFL contexts to enhance pronunciation. *TESL-EJ*, *11*(2), 1-11.

[12] Gilakjani, A. P. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education*, *1*(1), 1-15.

[13] Ministry of Education of the People's Republic of China. (2011). *China compulsory education primary school English curriculum standards* (2011 edition). Retrieved Nov. 5th, 2024, from http://www.moe.gov.cn/srcsite/A26/s8001/201112/t20111228_167340.html

[14] Huang, Y. C., & Liao, L. C. (2015). A Study of Text-to-Speech (TTS) in Children's English Learning. *Teaching English with Technology*, *15*(1), 14-30.

[15] Kong, N.Y. & Lin, J.S. (2001). Application of TTS whole-process speech technology in the making of CAI English multimedia Courseware. *China Electronic Education*, *(04)*, 45-47.

[16] Kondracki, N. L., Wellman, N. S., & Amundson, D. R. (2002). Content analysis: Review of methods and their applications in nutrition education. *Journal of nutrition education and behavior*, *34*(4), 224-230.

[17] Kilickaya, F. (2006). 'Text-To-Speech Technology': What Does It Offer to Foreign Language Learners? *CALL-EJ Online*, *7*(2), 7-2.

[18] Kataoka, H. (2007). *Preparation for university entrance examinations: How to learn frequently-appearing English words with the TTS audios*. A paper presented at the Multimedia & Internet Seminar in 2007, Japan Association for Language Education and Technology (LET), Kansai Chapter, held in Osaka on December 22, 2007.

[19] Kazazoğlu, S. (2013). Dictation as a language learning tool. *Procedia-Social and Behavioral Sciences*, *70*, 1338-1346.

[20] Li, H.Y. (2003). Using TTS Technology to Create English Situation of Audio-visual Integration [J]. *China Vocational and Technical Education*, *(13)*, 34-35.

[21] Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual review of applied linguistics*, *27*, 184-202.

[22] Nation, P. (1991). *Dictation, dicto-comp and related techniques*.

[23] Oller, J. W. (1971). Dictation as a Device for Testing Foreign-Language Proficiency1. *ELT Journal*, *25*(4), 254-259.

[24] Pellegrini, T., Costa, Â., & Trancoso, I. (2012). Less errors with TTS? A dictation experiment with foreign language learners. In *INTERSPEECH* (pp. 1291-1294).

[25] Seferoğlu, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology*, *36*(2), 303-316.

[26] Sha, G. (2010). Using TTS voices to develop audio materials for listening comprehension: A digital approach. *British Journal of Educational Technology*, *41*(4), 632-641.

[27] Sasirekha, D., & Chandra, E. (2012). Text to speech: a simple tutorial. *International Journal of Soft Computing and Engineering (IJSCE)*, *2*(1), 275-278.

[28] Valette, R.M. (1964). The use of the dictee in the French language classroom. *Modern Language Journal*, *48*(4), 431-434.

[29] Van Laere, E., & van Braak, J. (2017). Auditory support in linguistically diverse classrooms: factors related to bilingual text-to-speech use. *Journal of Computer Assisted Learning*, *33*(1), 79-95.

[30] Wen, Y. (2022). Application of intelligent speech system in primary school English teaching. *English for Campus*, *(09)*, 54-56.

[31] Whitaker, S.F. (1976). What is the status of dictation? *Audio-Visual Language Journal*, *14*(2), 87-93.

[32] Xu, J. & Dai, X.Y. (2010). Application of TTS technology in foreign language multimedia teaching. *Journal of Lanzhou Polytechnic College*, *(01)*, 31-34.

[33] Zhu, H. S. (2016). The design and application of TTS technology-assisted English listening. *Journal of Taiyuan City Vocational and Technical College*, *(12)*, 175-177.

[34] Zhang, J. (2020). *Design of elementary English listening teaching activity supported by TTS technology*. (Dissertation, Zhejiang Normal University).
[35] Zhu, X. (2023). A Study on the Application Model of Intelligent Text-to-speech Technology in Primary School English Listening Teaching. *English Teachers, (06)*, 97-100.

Suifang Yao was born in Guangzhou, China, on February 14, 1989. She graduated from Jilin University Zhuhai Campus in China with a bachelor's degree in Korean Language in 2012. She currently serves as an English teacher at Pui Ching Primary School in Guangzhou, China. Email: 190618217@qq.com

Kornwipa Poonpon was born in Nakornratchasima, Thailand. She holds a Ph.D. in Applied Linguistics from Northern Arizona University, Arizona, USA (2009). She holds a B.A. in English (First Class Honors) (1996) and an M.A. in Applied Linguistics from Mahidol University (2002), Thailand. She was the Head of the English Language Department, the Director of the Center for English Language Excellence, and the Chair of the M.A. in English Program at the Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand. Research interests include English language teaching, language assessment, and corpus linguistics. Assistant Professor Poonpon has also served as the President of the Thailand Association of Applied Linguistics (TAAL) (2023- 2024).