Journal of English Teaching
e-ISSN: 2622-4224 | p-ISSN: 2087-9628
http://ejournal.uki.ac.id/index.php/jet
Volume 10. Number 2, June 2024, pp. 183-197

The Effects of Reading Assistant Software on the Speech Fluency and Accuracy of EFL University Students

Nakhon Kitjaroonchai
Asia-Pacific International University
nakhon@apiu.edu

Suzanna Maywald
Asia-Pacific International University
suzanna@apiu.edu

Article History
Received: 5 April 2024
Accepted: 12 May 2024
Published: 12 June 2024

Keywords
EFL students, oral production, Reading Assistant Software, speech fluency, speech accuracy

Abstract
This study investigated the effects of a reading assistant software on the fluency and accuracy of speech production among EFL university students, employing a one-group pretest-posttest design. The participants included 50 Thai undergraduate English majors in their first and second years. Data were analyzed using paired-samples t-tests to measure L2 oral fluency, specifically pruned speech rate (PSR) and mean length of run (MLR). Accuracy rate was assessed by quantifying the proportion of error-free clauses within the participants' speeches from the pretest and posttest and calculating the average number of error-free clauses per T-unit. Results indicated significant improvements in speech fluency, evidenced by longer uninterrupted speech segments, fewer disfluencies like filled pauses and repairs, and increased overall fluency. Additionally, the post-test showed a higher percentage of error-free clauses compared to the pre-test, reflecting a substantial enhancement in grammatical accuracy. This suggests that the intervention had a positive impact on the participants' ability to produce grammatically accurate and error-free clauses. These outcomes suggest that the reading assistant software effectively enhances both the fluency and grammatical accuracy of speech production in EFL learners, underscoring its potential benefits in language education contexts.

How to cite this article (APA, 7th Ed.):
Kitjaroonchai, N. & Maywald, S. (2024). The effects of reading assistant software on the speech fluency and accuracy of EFL university students. *Journal of English Teaching, 10*(2), 183-197. https://doi.org/10.33541/jet.v10i2.5763

INTRODUCTION

It is widely recognized that English as a Foreign Language (EFL) learners possess a strong desire to achieve fluency in English within the language classroom. English has established itself as a lingua franca and a prominent language spoken globally (Crystal, 2003). However, it is evident that EFL learners exhibit significant variation in their oral performance in the target language (Natasia & Angelianawati, 2022). According to Elliott (1995), fluency and clear pronunciation are crucial elements for successful speech production. Numerous studies conducted with diverse samples of second language (L2) learners argue that individual differences in speech fluency may be influenced by various factors. These factors include cognitive maturity (Gathercole, 1999), frequency of exposure to the L2 (Paradis, Schneider, & Sorenson, 2013; Thordardottir & Brandeker, 2013), vocabulary size accumulated through learning experiences (Thorn & Gathercole, 1999; Thordardottir & Brandeker, 2013), as well as socioeconomic and first language (L1) backgrounds (Ellis Weismer et al., 2000; Sorenson, 2010). Despite the potential to develop fluency, many Thai students lack opportunities for extensive conversations in English. EFL classrooms in Thai schools focus more on memorizing grammar rules for exams rather than developing fluency, pronunciation, and communication skills (Arunsirot, 2017). As a result, Thai students struggle with extended conversations and English speaking proficiency (Sahatsathatsana, 2017).

After identifying English pronunciation difficulties faced by Thai students, Arunsirot (2017) used "Speech Analyzer" software to help Thai students improve their English pronunciation by comparing their speech to that of native speakers. This improved their accuracy. In 2020, Arunsirot found that augmented reality (AR) technology also enhanced pronunciation. Moxon (2021) studied speech recognition software's impact, showing positive effects on pronunciation. These studies suggest that reading aloud can enhance oral proficiency by adjusting pitch, tone, and volume, a finding supported by earlier research from Huang (2010) and Supraba, Wahyono, and Syukur (2020).

As technology enhances learning, higher education institutions are encouraged to integrate technological tools into language teaching (Wiwin, Utami & Taris, 2022). In response, the researchers' university acquired Reading Assistant (RA) software, a computer-based reading program with voice recognition and a digital library of over 300 books. While RA offers immediate feedback on mispronunciations, its potential to improve Thai EFL students' speaking performance remains unexplored. The researchers aim to investigate whether using RA software can enhance students' fluency and accuracy in speaking through oral reading.

LITERATURE REVIEW

L2 Speech Fluency

The definition of L2 speech fluency varies across studies in second language acquisition. It can include vocabulary range, grammatical correctness, pronunciation, idiomatic usage, and speech relevance, or specifically refer to the speed of oral production without hesitations (Lennon, 1990). L2 speech fluency is often understood as

the overall quality of oral proficiency (Wood, 2010). Skehan (2016) suggested assessing L2 fluency through measures like pause length and frequency, speech rates, phonation/time ratio, and self-corrections. Segalowitz (2016) proposed a framework with three domains: cognitive fluency (processing and converting ideas into speech), utterance fluency (flow of speech, articulation rate, mean length of uninterrupted speech), and perceived fluency (audience's evaluation of speech production). Utterance fluency, which is quantifiable, has received the most research attention.

Previous studies have examined utterance fluency by investigating various speech features. For instance, Bui and Huang (2016) conducted a study on L2 speech fluency and explored different aspects of fluency by using eight measures related to speed, including speed, stretch, voicing, mid-clause pauses, independent clause pauses, dependent clause pauses, filled pauses, and repairs. In the present study, we adopted the speed-related measures from Bui and Huang (2016) to assess cognitive and utterance fluency while engaging in L2 speaking activities.

L2 Speech Accuracy

Speech accuracy refers to how well a language learner produces language features according to the norms of a particular language, including syntax, lexis, and cohesion (Duijm, Schoonen, & Hulstijn, 2018). Duijm et al. (2018) note that speech accuracy is often evaluated alongside oral proficiency in language tests but can also be examined independently. Some researchers (e.g., Pallotti, 2009; Michel, 2017) argue that speech accuracy refers to error-free, native-like language use. Previous studies (e.g., Polio, 1997; Duijm et al., 2018) commonly use the word error rate (WER) method, which measures the percentage of incorrect transcriptions in speech using Automatic Speech Recognition (ASR) software. According to Errattahi, El Hannani, and Ouahmane (2018), the WER formula is defined as follows:

$$\text{WER} = (S + D + I)/N1$$

where I = total number of insertions, D = total number of deletions, S = total number of substitutions, and N1 = total number of input words.

An insertion (I) occurs when a word is added that was not uttered (for example, "BAT" becomes "battle"). A deletion (D) takes place when a word is left out of the transcript (for example, "take into consideration" becomes "consideration"). A substitution (S) happens when a word is altered (for example, "smoothly" is transcribed as "smoking")

While the Word Error Rate (WER) is commonly used, its limitations include lacking an upper bound and favoring insertions over deletions in noisy conditions (Erratahi et al., 2018). Polio (1997) used holistic scales, error-free clauses, and errors per 100 words. In this study, accuracy was evaluated by averaging error-free clauses per total clause (T-unit) and errors per 100 words, following Polio's (1997) approach.

Reading Assistant (RA) Software

The RA software, developed by Carnegie Learning, is part of the Fast ForWord program and provides real-time corrective feedback to speakers as they read aloud, helping them self-correct their pronunciation (Bhatt et al., 2020). This software has been shown to enhance learners' vocabulary acquisition, speech fluency, comprehension, and language prosody (Faisol et al., 2021; Li, 2020). It mimics a parent reading to a child, allowing learners to imitate sounds, intonation, rhythm, and prosody (Keller et al., 2018; Mahdi & Al Khateeb, 2019). The software automatically calculates the number of words read correctly per minute, helping instructors track reading competency and identify areas of weakness and strength (Bhatt et al., 2020). Keller et al. (2018) found that 195 out of 205 Thai EFL university students (95%) reported a positive experience using the software, noting significant improvements in their English language skills across reading, listening, and speaking.

Relationship between Reading and Speaking

Reading and speaking in a foreign language, such as English, are interrelated; improvement in one enhances the other (Novita, 2016; Li, 2020; Albadri & Halimah, 2022). Reading aloud aids comprehension and speaking skills, and extensive reading builds vocabulary, crucial for speaking performance (Anderson, 2018; Nation, 1994). Nunan (2003) highlights the role of reading aloud in developing speaking competence. Vocabulary acquired through reading significantly influences oral production (Li, 2020; Purwanto & Syafryadin, 2023). Novita (2016) suggests that extensive reading helps L2 learners become proficient speakers. Manurung (2015) found that integrating reading and speaking activities boosts self-confidence and vocabulary, enhancing oral presentations. Vocabulary knowledge is vital for reading comprehension in both L1 and L2 (Sidek & Rahim, 2015). Thus, vocabulary acquisition through reading is essential for speech production.

Although previous studies (e.g., Keller et al., 2018) have assessed learners' perceptions of the usefulness of the RA software, its impact on speech fluency and accuracy remains unexplored. This study aims to fill this gap by examining how the RA software affects Thai EFL university students' speaking fluency and accuracy. The study involves a minimum of 150 minutes of weekly usage of the RA software over ten weeks. The study is guided by the following research questions:

1. Can the RA software improve students' speech fluency?
2. Can the RA software enhance students' speech accuracy?

METHODOLOGY

Research Location and Participants

This quantitative study took place at a private international university in Saraburi province, Thailand. The researchers used purposive sampling to select 50 Thai undergraduate first- and second-year English majors, aiming for a sample with similar L1 backgrounds and English learning experiences. Participants, enrolled in reading courses between January and May 2023, were Thai speakers who had been learning English as an L2 for 9-13 years, averaging 11.5 years. Their English proficiency, based on TOEIC

scores, ranged from beginner (A1) to intermediate (B1) on the CEFR scale. The data collection was independent of their academic results or grades.

Instruments

Three instruments were employed to collect data: (1) Pre-Speaking Test, which involved a two-minute picture-description task, (2) Reading Assistant (RA) software, and (3) Post-Speaking Test, which also included a two-minute picture-description task.

Pre-Speaking Test

The pre-speaking test, held in researchers' offices during the first study week, involved a two-minute test using seven sets of picture sequences from online children's storybooks. Participants, who randomly selected a set, had two minutes to prepare and then creatively express a story or description. Their speeches were recorded with a computer voice recorder.

Reading Assistant (RA) software

The RA software is specifically designed to provide personalized education. It has the capability to listen to students as they read aloud and intervene when they mispronounce words. The software automatically scores students' oral reading performance and offers real-time guidance and feedback (Scientific Learning, 2021). In this study, the RA software was employed as an intervention. The participants were encouraged to log in and engage in daily oral reading practice for a minimum of 20-30 minutes per day (approximately 100 minutes per week), at their own convenience, over a period of ten weeks.

Post-Speaking Test

After the ten-week period of using the RA software, the post-speaking test was conducted in a manner similar to the pre-speaking test, taking place in the researchers' offices. Participants were instructed to randomly select a set of picture sequences and provide a two-minute speech, delivering a detailed story description. Their voices were recorded using the recorder system.

Data Collection

During the research phase, the coordinating researcher monitored participants' reading performance weekly using the RA software, tracking time spent and reading levels. The reading levels were based on grade equivalents provided by the software, ranging from first to second grade. For pre- and post-speaking tests, the researchers used VEED.IO automatic speech recognition (ASR) software to transcribe the speeches. The researchers reviewed these transcriptions to analyze speech fluency indicators, including articulation rate, short pauses, long pauses (>4 seconds), and mean length of run.

Data Analysis

The data collected from the pre- and post-speaking tests were transcribed primarily using the VEED.IO software. Subsequently, the researchers conducted a thorough review to calculate two key measures of L2 oral fluency: pruned speech rate (PSR) and mean length of run (MLR). These measures were adapted from the work of Bui and Huang (2016). The calculation of PSR and MLR was based on the formula outlined below:

$$PSR = \frac{(total\ words\ produced - vocal\ fillers - incomplete\ words - repeated\ word)\ x\ 60}{total\ time\ of\ speech\ (in\ second)}$$

$$MLR = \frac{\sum words\ in\ all\ stretches}{number\ of\ stretche}$$

In the study, an analysis of several variables was conducted, including: the number of long pauses (LP) lasting more than three seconds; the duration of pauses (DP); the frequency of filled pauses (FP), which refers to occurrences of vocalized sounds or words used as filler or hesitation markers during speech, such as 'um,' 'uh,' 'like,' 'you know,' or similar vocalizations; and the occurrence of repairs (RP), encompassing self-corrections, revisions, or substitutions of words or phrases.

Furthermore, in the present study, we assessed the accuracy rate by quantifying the proportion of error-free clauses within the participants' two-minute speeches. Additionally, we calculated the average number of error-free clauses per T-unit. The analysis of these accuracy measures allowed us to align our findings with the definitions provided in Table 1, as presented below.

Table 1:
Definition of terms and calculation for accuracy rate

Terms	Definitions
Total clauses (T-units)	An independent clause and any clauses dependent on it
Percentage of error-free clauses	Percentage of clauses which do not contain any error to the total number of clauses
Errors per 100 words	Number of errors divided by the total number of words produced divided by 100

Below is an example of analyzing T-units, percentage of error free clauses, and errors per 100 words.

> My name is Santa and I want to describe these pictures. It's at a park, in a center of the city, on the heavy rainy day. There is a stray little dog at the park. There are two ladies walk into this park. So they see the little dog and they took pity at that dog, so they take a dog to their house. After that a few days, they go out around the city to announce," who is the owned a dog?" But no one said the dog was theirs and so the two ladies decided to adopt this dog. With a kind cuteness of this dog, so they adopted it. The two ladies then gave some milk to this dog, and sometimes they play with a dog and don't let the dog feel lonely anymore. They gave the name to this dog. And, the name of the dog is Mona, it's very cute name. So the dog is now happy to be adopted, and the two ladies love this dog like their own sister.

The breakdown of the paragraph into T-units:

1. My name is Santa and I want to describe these pictures.
2. It's at a park, in a center of the city, on a heavy rainy day.
3. There is a stray little dog at the park.
4. There are two ladies walk into this park.
5. So they see the little dog and they took pity at that dog, so they take a dog to their house.
6. After that a few days, they go out around the city to announce," who is the owned a dog?"
7. But no one said the dog was theirs and so the two ladies decided to adopt this dog.
8. With a kind cuteness this dog, so they adopted it.
9. The two ladies then give some milk to this dog, and sometimes they play with a dog and don't let the dog feel lonely anymore.
10. They gave the name to this dog.
11. And, the name of the dog is Mona, it's very cute name.
12. So the dog is now happy to be adopted, and the two ladies love this dog like their own sister.

The paragraph contains 12 T-units. Each T-unit represents a complete thought or idea that can stand on its own or can be combined with other T-units to form a coherent paragraph or text.

Out of these 12 T-units, six of them are grammatically correct: T-unit items 1, 3, 7, 9, 10, and 12. However, the underlined T-units contain grammatical errors, which can be corrected as follows:

2. "It's in a park, in the center of the city, on a heavy rainy day."
4. "There are two ladies walking into this park."
5. "So they see the little dog and they take pity on that dog, so they take the dog to their house."
6. "After that a few days, they go out around the city to announce, 'Who owns a dog?"
8. "With the dog's kind cuteness, they adopted it."
11. "And the name of the dog is Mona; **it's a** very cute name."

The paragraph consists of 176 words and contains 18 errors. Examples of word errors have been identified using underlines, and the corrections have been made as follows.

My name is Santa and I want to describe these pictures. It's **at** [in] a park, in **a** [the] center of the city, on **the** [a] heavy rainy day. There is a stray little dog at the park. There are two ladies **walk** [walking] into this park. So they see the little dog and they **took pity at** [take pity on] that dog, so they take **a** [the] dog to their house. After that a few days, they go out around the city to announce," who **is the owned** [owns] a dog?" But no one said the dog **was** [is] theirs and so the two ladies decided to adopt this dog. With **a kind cuteness of this dog**, [the dog's kind cuteness] so they adopted it. The two ladies then give some milk to this dog, and sometimes they play with **a** [the] dog and don't let the dog feel lonely anymore. They **gave** [give] the name to this dog. And, the name of the dog is Mona, **it's very** [a] cute name. So the dog is now happy to be adopted, and the two ladies love this dog like their own sister.

From the analysis, we may conclude that error per 100 words = 18/176 x 100 = 10.22

FINDINGS

For Research Question 1, examining RA software's impact on speech fluency, a paired-samples t-test was employed to analyze the pre- and post-speaking test mean scores. Key elements included word counts per minute, pruned speech run, mean length of run, long pauses, pause duration, filled pauses, and repairs. Results are presented in Table 2.

Table 2: *Mean comparison of pre- and post-speaking test: speech fluency test*

	Variables	N	M	SD	MD	*t*	df	Sig. (2-tailed)
Pair 1	WCPre	50	172.02	45.26	-5.28	-1.22	49	0.228
	WCPost	50	177.3	45.85				
Pair 2	PSRPre	50	86.01	22.63	-2.64	-1.22	49	0.228
	PSRPost	50	88.65	22.92				
Pair 3	MLRPre	50	8.93	2.61	-2.08	-6.29	49	0.000
	MLRPost	50	11.01	1.87				
Pair 4	LPPre	50	0.82	1.22	0.64	3.88	49	0.000
	LPPost	50	0.18	0.39				
Pair 5	DPPre	50	2.76	4.20	2.06	3.58	49	0.001
	DPPost	50	0.70	1.52				
Pair 6	FPPre	50	11.34	9.27	4.72	5.29	49	0.000
	FPPost	50	6.62	5.20				
Pair 7	RPPre	50	10.46	5.49	4.54	7.77	49	0.000
	RPPost	50	5.92	3.83				

Note: WCPre = Word Counts Pre-Speaking Test; WCPost = Word Counts Post-Speaking Test PSRPre =Pruned Speech Run Pre-Speaking Test; PSRPost = Pruned Speech Run Post-Speaking Test; MLRPre = Mean Length of Run Pre-Speaking Test; MLRPost= Mean Length of Run Pre-Speaking Test; LPPre = Number of Long Pauses Pre-Speaking Test; LPPost= Number of Long Pauses Post-Speaking Test' DPPre = Duration of Pause Pre-Speaking Test; DPPost = During of Pause Post-Speaking Test; FPPre = Number of Filled Pauses Pre-Speaking Test; FPPost = Number of Filled Pauses Post-Speaking Test; RPPre= Number of Repairs Pre-Speaking Test; RPPost= Number of Repairs Post-Speaking Test; MD = Mean Differences

As shown in Table 2, the findings indicate significant mean differences on MLR with $t(49) = -6.29$, $p < .001$. Results show that the mean score for MLR in the pre-test was (M = 8.93 (SD = 2.61), while the mean score for MLR in the post-test was (M = 11.01, SD = 1.87). This suggests that, on average, participants produced longer runs of speech during the post-test compared to the pre-test. The findings also indicate significant mean differences on number of LP with $t(49) = 3.88$, $p < .001$. Results show that the mean score for LP in the pre-test was (M = 0.82, SD = 1.22), while in the post-test, the mean score was (M = 0.18, SD = 0.39). This suggests that, on average, participants exhibited a lower frequency of LP during the post-test compared to the pre-test. This was further supported by a decrease in the duration of pauses (DP) with $t(49) = 3.58$, $p = 0.001$ from the pre-test to the post-test. The mean score for DP in the pre-test was (M =2.76, SD = 4.20), while mean score for DP in the post-test decreased to (M = 0.70, SD = 1.52). This suggests that, on average, participants had shorter pauses during the post-test compared to the pre-test. The decrease in the mean score indicates an improvement in participants' ability to maintain a smoother and more fluent speech production, as they demonstrated a reduced duration of pauses. The results also show significant mean differences on FP with $t(49) = 5.29$, $p < .001$. Results indicate that the mean score for FP in the pre-test was (M =11.34, SD = 9.27), while in the post-test, the mean score decreased to (M =6.62, SD = 5.20). This suggests that, on average, participants used fewer filled pauses during the post-test compared to the pre-test. Overall, these findings suggest that an enhancement in participants' ability to manage their pauses and maintain a more fluent and uninterrupted speech production. Furthermore, the findings show significant mean differences on RP with $t(49) = 7.77$, $p <.001$. Results show a decrease in RP from the pre-test to the post-test. The mean score for RP in the pre-test was (M =10.46, SD = 5.49), while in the post-test, the mean score decreased to (M = 5.92, SD = 3.83). This suggests that, on average, participants made fewer repairs during the post-test compared to the pre-test.

From the results of the study, we noticed evidence of improvements in speech fluency. Participants demonstrated enhanced performance in terms of longer uninterrupted speech segments, reduced occurrence of disfluencies such as filled pauses and repairs, and increased fluency in their speech production. These findings indicate an overall enhancement in participants' fluency skills, as evidenced by their ability to maintain smoother and more cohesive speech delivery. The decrease in variability observed in several measures further suggests a more consistent performance among participants in the post-test. These results highlight the effectiveness of the intervention in promoting fluency in speech production.

For Research Question 2, examining RA software's impact on speech accuracy, a paired-samples t-test was used to analyze the pre- and post-speaking test scores. Elements included T-units, error-free clauses percentage, and errors per 100 words. Results are shown in Table 3, providing a comprehensive overview of the statistical outcomes.

Table 3. *Mean comparison of pre- and post-speaking test: speech accuracy test*

Variables		N	M	SD	MD	*t*	df	Sig. (2-tailed)
Pair 1	T-unitPre	50	10.48	2.31	0.12	0.40	49	0.686
	T-unitPost	50	10.36	2.06				
Pair 2	PEFCPre	50	23.86	8.23	-14.34	-13.47	49	0.000
	PEFCPost	50	38.20	9.55				
Pair 3	EPWPre	50	19.57	7.13	8.03	10.17	49	0.000
	EPWPost	50	11.53	2.99				

Note: T-unitPre = T-unit Pre-Speaking Test; T-unitPost = T-unit Post-Speaking Test; PEFCPre= Percentage of Error Free Clauses Pre-Speaking Test; PEFCPost = Percentage of Error Free Clauses Post-Speaking Test; EPWPre = Error Per 100 Words Pre-Speaking Test; EPWPost= Error Per 100 Words Post-Speaking Test MD = Mean Differences

As shown in Table 3, the findings of the paired samples *t*-test indicate that there is no significant difference between the pre- and post-test scores for the T-unit measure with $t(49) = 0.40$, p = .686, which is greater than the commonly used significance level of 0.05. The mean score for the T-unit measure in the pre-test was (M = 10.48, SD = 2.31), while the mean score in the post-test was (M = 10.36, SD = 2.06). Based on these results, we can conclude that there is no evidence of a significant change in the T-unit scores between the pre- and post-test conditions. However, the findings indicate a significant difference in PEFC between the pre-test and post-test with $t(49) = -13.47$, $p < .000$, indicating an extremely small probability of obtaining such results by chance. The mean score for PEFC in the pre-test was (M = 23.86, SD = 8.23), while the mean score in the post-test was (M = 38.20, SD = 9.55). The mean difference between the pre- and post-test scores was -14.34. Based on these results, we can conclude that there is a significant improvement in PEFC from the pre-test to the post-test. Participants demonstrated a substantial increase in the production of error-free clauses, with the post-test mean exceeding the pre-test mean by 14.34 units. The negative mean difference indicates that, on average, participants produced a higher percentage of error-free clauses in the post-test compared to the pretest.

The findings of the paired samples *t*-test further indicate a significant difference in EPW between the pre-test and post-test with $t(49) = 10.17$, $p < .000$. The mean score for EPW in the pre-test was (M = 19.57, SD = 7.13), while the mean score in the post-test was (M = 11.53, SD = 2.99). The mean difference between the pre- and post-test scores was 8.03. Based on these results, we can conclude that there is a significant improvement in EPW from the pre-test to the post-test. Participants demonstrated a notable decrease in the number of errors in their speech production, with the post-test mean being lower than the pre-test mean by 8.03 units. The positive mean difference indicates that, on average, participants made fewer errors per 100 words in the post-speaking test compared to the pre- speaking test. This improvement suggests that the intervention or treatment implemented between the two tests had a positive impact on participants' ability to reduce errors in their speech production.

The results of the paired samples *t*-tests revealed interesting findings regarding the pre-and post-test measures. Firstly, there was no significant difference in the T-unit measure between the two tests, indicating that participants' performance in terms of T-units remained relatively stable. However, a significant improvement was observed in the PEFC measure, indicating that participants produced a higher percentage of error-free clauses in the post-test compared to the pre-test. Additionally, there was a significant improvement in EPW measure, indicating that participants made fewer errors in their speech production during the post-test compared to the pre-test. These findings highlight the positive effects of the intervention on participants' linguistic performance, specifically in terms of producing error-free clauses and reducing errors per 100 words.

DISCUSSION

The RA Software Helps Improve Students' Speech Fluency

Speech fluency is essential for effective communication, and improving fluency is a common goal (Lennon, 1990; Bui & Huang, 2016; Moxon, 2021). The study showed significant improvements in speech production, including longer uninterrupted speech, fewer disfluencies, and increased overall fluency, highlighting the effectiveness of RA software. Participants exhibited longer speech segments during the post-speaking test, suggesting enhanced fluency, likely due to the advanced speech verification technology and scientifically based interventions of the RA software (Scientific Learning, 2021). The software aids students by strengthening vocabulary and comprehension, facilitating self-assessment, and improving fluency (Bhatt et al., 2020; Faisol et al., 2021). These findings align with Arunsirot (2020), who found that augmented reality technology significantly improved college students' pronunciation and comprehension skills.

The findings reveal a significant decrease in the frequency and duration of long pauses (LP) and pauses (DP) between the pre- and post-test conditions. Participants exhibited fewer and shorter pauses in the post-test, indicating improved speech fluency. This progress may be attributed to the extensive reading content provided by the RA software, which promotes consistent practice and nurtures oral fluency. Routine reading enhances vocabulary (Huang, 2010; Wood, 2010; Faisol et al., 2021), offering exposure to diverse words and contexts. This increased lexical knowledge facilitates smoother transitions between ideas and reduces the need for pauses or hesitations, enabling more precise and effective self-expression (Segalowitz, 2010, 2016; Novita, 2016).

The RA Software Enhances Students' Speech Accuracy

Speech accuracy is crucial for L2 oral production as it impacts communication effectiveness (Manurung, 2015; Bui & Huang, 2016; Arunsirot, 2017). Accurate pronunciation, grammar, and vocabulary help learners convey messages clearly and build confidence and credibility, fostering positive communication (Wood, 2010; Arunsirot, 2020; Supraba et al., 2020). This study focused on three elements of speech accuracy: T-units, percentage of error-free clauses (PEFC), and errors per 100 words (EPW). The results showed no significant difference in T-unit measures between the pre- and post-

tests, indicating stable performance. However, there was a significant increase in PEFC, suggesting that the intervention positively impacted participants' ability to produce grammatically accurate clauses. These findings align with Li's (2020) study, which found that the Speech Recognition System (SRS) improved English pronunciation and accurate language use.

There was a significant decrease in errors per 100 words (EPW) between the pre- and post-speaking tests, with a mean difference of 8.03. This indicates fewer errors in post-test speech production, demonstrating the RA software's positive impact on participants' proficiency. The intervention led to more grammatically accurate and error-free speech, enhancing participants' ability to express thoughts and ideas. These findings align with previous research (e.g., Mahdi & Al Khateeb, 2019), which shows that RA software motivates learners to improve pronunciation, language proficiency, and fosters a positive attitude towards self-directed learning. The results underscore the effectiveness of RA software in reducing speech errors and enhancing overall communication skills.

IMPLICATIONS

The present study provides some implications for teaching oral production by integrating the RA software to improve speech fluency and accuracy.

Integration of RA Software for Improving Speech Fluency

The research demonstrates that the RA software can improve students' speech fluency, which is vital for effective communication. Educators should integrate the software into their teaching methods, as it uses advanced speech verification and scientific interventions to reduce disfluencies and increase fluency. Regular use of the software aids vocabulary acquisition and smoother speech transitions, reducing hesitation. Therefore, incorporating the RA software can enhance students' speaking skills and overall communication effectiveness.

Focused Language Interventions for Speech Accuracy

The research shows that the RA software improves students' speech accuracy. Educators should use focused interventions to enhance grammatical accuracy and reduce errors. Incorporating speech recognition systems can promote error-free clauses and fewer errors per 100 words, enhancing language proficiency. By offering targeted practice and feedback educators can help learners improve their pronunciation, grammar, and vocabulary, leading to clearer communication.

CONCLUSION

The research highlights the significant positive impact of integrating RA software on students' speech fluency and accuracy in oral production. The intervention led to substantial improvements in speech production, including longer uninterrupted speech, reduced disfluencies, and increased overall fluency. The software's advanced speech verification technology and scientifically based interventions supported vocabulary and

comprehension skills, contributing to enhanced fluency. Moreover, the RA software positively influenced students' speech accuracy, with notable improvements in error-free clauses and reduced errors per 100 words, indicating enhanced grammatical accuracy. These findings underscore the pedagogical implications of integrating RA software to promote effective communication and language proficiency. Educators are encouraged to consider RA software as a valuable tool for enhancing students' speaking abilities and overall language development in EFL or ESL classroom settings

ACKNOWLEGMENTS

We would like to thank the Research, Ranking and Development Committee of Asia-Pacific International University for providing financial support for this project (RRDC 2022-155).

REFERENCES

Albadri, A., & Halimah, N. (2022). The correlation between reading and speaking skills on students' English learning as foreign language. *JOEY: Journal of English Ibrahimy*, *1*(2), 27-42.

Anderson, L. W. (2018). A critique of grading: Policies, practices, and technical matters. *Education Policy Analysis Archives, 26*, 1–31. https://doi.org/10.14507/epaa.26.3814

Arunsirot, S. (2017). Implementing a speech analyzer software to enhance English pronunciation competence of Thai students. *Journal of Education Studies, Burapha University*, *28*(2), 116-129.

Arunsirot, N. (2020). Implementing the augmented reality technology to enhance English pronunciation of Thai EFL students. *KKU Research Journal (Graduate Studies) Humanities and Social Sciences*, *8*(3), 142-153.

Bhatt, C., Kothari, P., Masrani, S., & Kokel, V. (2020). Reading Assistant: a reciter in your pocket. In *Wearable and Implantable Medical Devices* (pp. 163-178). Academic Press.

Bui, G., & Huang, Z. (2016). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research*, *22*(1), 94-114.

Crystal, D. (2003). *English as a global language*. Cambridge University Press.

Duijm, K., Schoonen, R., & Hulstijn, J. H. (2018). Professional and non-professional raters' responsiveness to fluency and accuracy in L2 speech: An experimental approach. *Language Testing*, *35*(4), 501-527.

Elliott, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal*, *79*(4), 530-542.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-aged children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 865–878.

Errattahi, R., El Hannani, A., & Ouahmane, H. (2018). Automatic speech recognition errors detection and correction: A review. *Procedia Computer Science*, *128*, 32-37.

Faisol, M. A. M., Ramlan, S. A., Saod, A. H. M., Mozi, A. M., & Zakaria, F. F. (2021). Mobile-based speech recognition for early reading assistant. In *Journal of Physics: Conference Series* (Vol. 1962, No. 1, p. 012044). IOP Publishing.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*, 410–419.

Huang, L. (2010). Reading aloud in the foreign language teaching. *Asian Social Science Journal, 6*(4), 148-150.

Keller, K., Keough, W., & Galgao, F. (2018, October). Student response to use of reading assistant software for English language learners in Thailand. In *Abstract Proceedings International Scholars Conference* (Vol. 6, No. 1, pp. 284-284).

Lennon, F. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-417.

Li, J. (2020). An empirical study on reading aloud and learning English by the use of the reading assistant SRS. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(21), 103-117.

Mahdi, H. S., & Al Khateeb, A. A. (2019). The effectiveness of computer-assisted pronunciation training: A meta- analysis. *Review of Education*, *7*(3), 733-753.

Manurung, K. (2015). Improving the speaking skill using reading contextual internet-based instructional materials in an EFL class in Indonesia. *Procedia-Social and Behavioral Sciences*, *176*, 44-51.

Michel, M. (2017) Complexity, accuracy and fluency (CAF). In Shawn Loewen & Masatoshi Sato *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge.

Moxon, S. (2021). Exploring the effects of automated pronunciation evaluation on L2 students in Thailand. *IAFOR Journal of Education*, *9*(3), 41-56.

Natasia, G., & Angelianawati, L. (2022). Students' perception of using storytelling technique to improve speaking performance at SMPN 143 Jakarta Utara. *Journal of English Teaching*, *8*(2), 282-292.

Nation, I. S. P. (Ed.). (1994). New ways in teaching vocabulary. Alexandria, VA: TESOL.

Novita, L. (2016). Relationship between reading and pronunciation and students' speaking skills. In *Proceeding of the International Conference on Teacher Training and Education* (Vol. 2, No. 1, pp. 527-532).

Nunan, D. (2003). *Practical English Language Teaching*. Boston: McGraw Hill.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*,30(4), 590-601.

Paradis, J., Schneider, P., & Sorenson Duncan, T. (2013). Discriminating children with language impairment among English-language learners from diverse first-language backgrounds. *Journal of Speech, Language, and Hearing Research, 56*, 971–981.

Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*, 101–143.

Purwanto, A. A., & Syafryadin. (2023). Students' perception on using Duolingo for learning English vocabulary. *JET (Journal of English Teaching)*, *9*(1), 70-82.

Sahatsathatsana, S. (2017). Pronunciation problems of Thai students learning English phonetics: A case study at Kalasin University. *Journal of Education, Mahasarakham University*, *11*(4), 67-84.

Scientific Learning (2021). An online guided reading tool for improving vocabulary, fluency, comprehension, and prosody. Retrieved from https://www.scilearn.com/products/fast-forword/ reading-assistant

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York, NY: Routledge.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics, 54*(2), 79-95

Sidek, H. M., & Rahim, H. A. (2015). The role of vocabulary knowledge in reading comprehension: A cross-linguistic study. *Procedia-Social and Behavioral Sciences*, *197*, 50-56.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics*, *36*, 34-49.

Sorenson Duncan, T. (2010). *Emerging L2 consonant systems and performance on non word repetition tasks* (Master's thesis). University of Alberta, Edmonton, Alberta, Canada.

Supraba, A., Wahyono, E., & Syukur, A. (2020). The implementation of reading aloud in developing students' speaking skill. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, *8*(1), 145-153.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46*, 1–16.

Thorn, A., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non- bilingual children. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 52*(A), 303–324.

Wiwin, D., Utami, U. W., & Taris, T. (2022). Digital media and its implication in promoting students' autonomous learning. *JET (Journal of English Teaching)*, *8*(1), 97-106.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom application*. London: Continuum.