【Research Note】

A Feasibility Study on Audio-assisted Extensive Reading and Reading Rate Growth

Cheryl Kirchhoff, Miguel Mision, Jean-Pierre J. Richard

Introduction

In second language learning contexts, developing reading fluency is an important objective in itself, and a means to overall language improvement. Extensive reading is one teaching practice which can develop reading fluency as shown in reading rate improvement (Beglar, Hunt & Kite, 2012). In addition, reading while listening to audio narration offers a more comprehensive experience with a text than silent reading. Audio-assisted reading gives the learner correct pronunciation of new words, prosody awareness, voice differentiation of multiple characters, and background sounds. If audio-assisted reading scaffolds learners' reading skills and comprehension, would it also increase reading rate? Note that in this paper, reading rate refers solely to the number of words that learners read per minute, in other words, how quickly one reads. In contrast, reading fluency is a broader term which encompasses reading rate, accuracy, comprehension, and perhaps prosody if reading aloud. Audio-assisted extensive reading and its impact on reading fluency has been the subject of several studies (Chang & Millet, 2015; Gobel, 2011; Milliner, 2019). This paper explains an attempt to capture the influence of audio-assisted reading on reading rate within an extensive reading program. Weaknesses in the method of measuring reading rate yielded inconclusive results, thus plans to improve the research are included.

Literature Review

Reading fluency is the ability to read quickly with comprehension. Fluency

includes multiple skills such as rapid decoding, increasing the size of automatically recognized chunks, extensive exposure to text, accuracy in comprehension, and time for reading skills to develop (Grabe, 2010, Nation & Waring, 2020). A teaching practice that supports reading fluency is extensive reading in which learners read very large amounts of text that the learner can understand. Studies that compared reading fluency between learners who did extensive reading and those who did intensive reading, have found extensive reading to result in greater fluency (Beglar, Hunt & Kite, 2012; Huffman, 2014; McLean & Rouault, 2017; Park, 2017). Extensive reading originally utilized paper texts, but now is often done in online libraries which include an audio narration function.

Adding audio narration to reading can scaffold learners' comprehension of text. Audio narration adds pronunciation, pauses that signal phrases of meaning, and sound effects that support learners' comprehension and concentration. Audio-assisted extensive reading has been studied for its influence on reading fluency and other variables. A study of Japanese university students that did audio-assisted reading of 10 graded readers over 10 months reported increases in TOEFL IP scores; however, other factors, including learning in other English courses were not controlled and may have also influenced the results (Gobel, 2011). Despite the reported amount of audio-assisted reading not being large; Gobel observed that the amount of reading correlated with gains in the TOEFL IP scores in this study and a following study. A study that compared audio-assisted extensive reading to silent reading found greater increases in reading and comprehension for the audio-assisted extensive reading group (Chang & Millet, 2015). The Taiwanese high school students in the study experienced a highly controlled version of extensive reading in which all students read the same book in class and talked about the stories with classmates. In the year-long program students read 20 graded readers. Chang and Millet suggested that the audio helped learners concentrate on reading. However, an important limitation of the study by Chang and Millet is that learners had little agency; the lessons, the materials, and the pacing were all highly controlled by the teacher. A similar study comparing audio-assisted extensive reading and silent reading was conducted at a Japanese university in which students self-selected their books (Milliner, 2019). The 15-week study also found that the audio-assisted group improved reading and listening skills more than the silent reading group. Several limitations with the study by Milliner are that the research groups (control, treatment 1, treatment 2) were small (17-22 students per group), there is no information about initial TOEIC scores, and importantly any TOEIC "gains" were likely within the

standard error of difference, which poses the question: Where any gains achieved at all?

Requesting students to do audio-assisted extensive reading outside the classroom requires preparation to ensure students choose to listen while reading. Kirchhoff and Mision (2022) explored students' willingness to listen to audio while reading and their opinion of the two-modal approach. In the middle of an 8-month extensive reading program, the instructors explained the value of listening while reading, demonstrated how to use audio functions, and asked students to do audio-assisted reading for the first half of a seven-week term. After experiencing audio-assisted reading most students reported that it was good English practice, and many students reported being able to concentrate or felt absorbed in their reading. Students also explained reasons for not listening while reading: listening requires having earphones and an appropriate environment, and following along with the audio requires concentration which can be exhausting.

After learning from students' experiences and opinions of audio-assisted reading, the writers wanted to know if adding a second mode of input (listening) would influence students' reading rates. Reading rates would be measured before and after the audio-assisted assignment. We hypothesized that dual modality input would scaffold comprehension and lead learners to read at a faster pace with comprehension. Thus, the research questions were,

RQ 1: Did reading rate change from a pre-test to a post-test and then a delayed post-test?

RQ 2: Does audio-assisted extensive reading lead to reading rate gains?

Methodology

Context

This study took place at a small, public university in central Japan. Students at this university are non-English majors, who complete a required two-year English program in the first two years of university. The university uses a quarter system, with each quarter comprised of seven weeks of classes and approximately one week of exams. The first two quarters are equivalent to the spring (first) semester in Japanese higher education, and the last two quarters are equivalent to the fall (second) semester. Students have four 100-minute lessons per week in Year 1, and two-to-four 100-minute lessons per week in Year 2. Half of the lessons are taught by Japanese faculty members, focusing on English accuracy; and half of the classes are taught by non-Japanese faculty, focusing on English fluency. As part of the first-

year common curriculum, students are introduced to extensive reading in the first quarter, freely using an online extensive reading library (Xreading.com), to read three-to-five books, and participate in several L2-English book talks (i.e., short book reports in pairs and groups) about the books they are reading. From the second quarter, Xreading is used for accountability - students must read 20,000 words in each of the halves of Quarter 2, and then 30,000 words in each of the halves of Quarters 3 and 4. In all, they are required to read 160,000 words between June and January. For each half of Quarters 2-4 in which students reach class targets, they receive up to five points (i.e., up to 10 class points, 10% of final grade) are awarded to students in Quarters 2-4). In 2022, students read 141,000 words on average. Reading is primarily done outside of class.

In addition to the extensive reading program, the English courses with an emphasis on fluency include learning basic communication skills and discussion. English courses also require students to study the New General Service List (NGSL). Students take a levels test to gauge their vocabulary comprehension relative to the NGSL, and at the end of each quarter are given an NGSL final exam based on their level - worth 10% of their total assessment score. Students who achieve a score of 80% or more 'pass' the exam and move up to the next level of the NGSL.

Participants

The study took place in eight of 10 first-year fluency classes in Quarter 3 (i.e., late September to mid-November) of 2022, and initially 206 students were potential participants. Care was taken to include only students who did extensive reading and reached 60% of comprehension questions on reading rate measurements. This percentage was chosen because it is the minimum score required to show evidence of having read a book in Xreading. Extensive reading was defined as reading 60,000 words, plus or minus 15,000. Outliers, those who read less than 45,000 words, and those who read more than 75,000, were deleted. This eliminated two advanced extensive readers and 28 non-extensive readers. Subsequently, students who scored 5 points or less on 10 comprehension questions were removed as they did not show that they comprehended the reading. This was 19 students on the Pre-Test, a further 20 students on the post test, and 14 students on the delayed post test. The remaining students (n = 125) completed all tests, read between 45,000-75,000 words, and generally comprehended the readings. On average, these students read 60,663.9 words (SD = 4,838.4), and average scores on the 10 comprehension questions across the three tests were high, 8.2 (SD = 1.7) 8.4 (SD = 1.0), and 8.1 (SD = 1.0). This group

included three international students from other Asian settings, and a small number of bi-cultural Japanese students. These students were retained.

Design

In Class 1 of Quarter 3, students were told the purpose of the research, and had two practice readings with comprehension questions to measure reading rate (*i.e.*, reading rate measure, RRM). Instructors projected a stopwatch onto the class screen. Students read the first reading, recorded their time, then continued to the comprehension questions. After students had finished the questions for the first reading, these were checked, and students noted their results in the test booklet. This process was repeated for the second reading of the practice RRM. In Class 2, students completed the pre-RRM that was used as the starting point. The procedures were the same as for the practice in Class 1. In Class 13, approximately six weeks after the pre-RRM, students completed the post-RRM. In Class 9 of Quarter 4, approximately nine weeks after the pre-RRM, students completed the delayed post-RRM. The procedures for the post-RRM and delayed post-RRM were the same as above. See Figure 1 for clarification.

Quarter	3 (late September to mid-November)								4 (late November to late January)							
Class	1	2	…	7	8	…	13	14	1	2	…	7	8	9	…	14
XR	30,000 words				30,000 words				30,000 words				30,000 words			
RRM	0	1					2							3		
Tests																
RWL	Required				Optional											

Note. RRM tests are 0 = practice, 1 = pretest, 2 = posttest, and 3 = delayed posttest.

Figure 1. Timeline for the AER Research Project.

Materials

At the start of the third quarter, students were given an assignment to view an instructor-made web page that explained audio-assisted reading (https://sway.office.com/pLrz964i1Ud04jkt?ref=Link). The materials included Japanese (L1) video and written explanations of audio-assisted reading. Also, brief Japanese summaries of English research on audio-assisted extensive reading explained benefits of audio assisted reading (Chang, 2011; Woodall, 2010). Student comments about audio-assisted reading from Campbell, Calman, and Campbell's research (2021) were also included. In addition, a video demonstrated how to use the Xreading audio controls,

including audio speed adjustment. Students were required to complete questions to demonstrate that they understood the materials.

The Xreading virtual library has over 1200 books for English language learners. The audio narration that accompanies each digital book is typically a single narrator who slightly varies their voice for different characters. Some of the narrations have background noises for artistic effect. The most common dialects spoken by the readers are British or North American English dialect. Each page of a book has its own audio file. The audio files have 5 speed levels, two that are slower than normal and two that are faster than normal.

The third quarter reading assignment for the first half (26 days) was to read 30,000 words while listening to the audio narration. The assignment for the second half of the quarter (26 days) was also 30,000 words; however, students could choose to read-only or listen to audio while reading. Each assignment was 5% of the final grade for the quarter.

All reading texts for the reading rate measurements were chosen from Nation and Malarcher's "Reading for Speed and Accuracy" (2007). Topics from the textbook were chosen based on their familiarity to students. Length of texts was evaluated by standard word count which includes punctuation and spaces of text (Nation & Waring, 2020). Care was taken to ensure all texts had similar standard word counts (see Appendix for list of readings). The vocabulary from each text was analyzed using the Complete Lexical Tutor (https://www.lextutor.ca/). Thus, texts were chosen based on the highest percentages of the first 2000 NGSL words – most researchers agree that the minimum lexical coverage for sufficient L2 reading comprehension is the point at which the reader can comprehend at least 95% of the vocabulary (Schmitt, Jiang & Grabe, 2011). Finally, off-list words (words that did not appear in the NGSL or NAWL lists) were also taken into account because some crucial words being misunderstood by a majority of students would jeopardize the overall comprehension of the text. Thus, the reading rate measurement texts were thought to be similar in content, word count and vocabulary level.

Data Analysis

Data were analyzed using JASP, Version 0.17.1 (JASP Team, 2023). First, descriptive statistics were gathered. Second, for RQ1, two paired-sample student *t*-tests were run to test whether words per minute (WPM) at Time 1 (*i.e.*, pretest) were different from at Time 2 (*i.e.*, posttest) and Time 3 (*i.e.*, delayed posttest). Using the Bonferroni correction, the *p*-value significance was set to $p = 0.025$ (0.05/2). For

RQ2, a simple linear regression, with bootstrapping, was used to investigate the relationship between total AER and variance in WPM growth.

Results and Discussion

Table 1 displays descriptive statistics. Students read on average 60,304 words, for 8.5 hours (30,680.9/3600). Listening time was approximately half as long, 4.8 hours (17189.8/3600). WPM appears to have increased between Time 1 (Pre-test) and Time 2 (Posttest); however, WPM fell to Pre-test levels at Time 3 (Delayed posttest). Except for Total Listening time, the other variables were highly skewed and kurtotic. For example, see Figure 2, WPM counts for the pre-test.

Table 1. Descriptives for Average Total Words, Reading and Listening Time, and WPM (n = 125)

	Total Words	Total Reading Time (in seconds)	Total Listening time (in seconds)	Pre-Test WPM	Posttest WPM	Delayed Posttest WPM	WPM Growth between pre- and posttests.
M	60303.7	30680.9	17189.8	132.7	159.1	132.1	27.3
SD	4685.3	8216.6	8872.2	43.0	40.8	36.0	19.6
Skew	−1.3	1.1	−0.3	1.4	1.4	1.6	1.0
SE of Skew	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Kurtosis	3.8	2.2	−0.8	2.9	2.6	3.8	0.3
SE Kurtosis	0.4	0.4	0.4	0.4	0.4	0.4	0.4
S-W	0.8	0.9	1.0	0.9	0.9	0.9	0.9
S-W p	<.001	<.001	0.004	<.001	<.001	<.001	<.001
Min	45301.0	14980.0	0.0	70.2	89.0	81.1	0.0
Max	72853.0	60960.0	34100.0	307.1	307.1	293.5	83.3

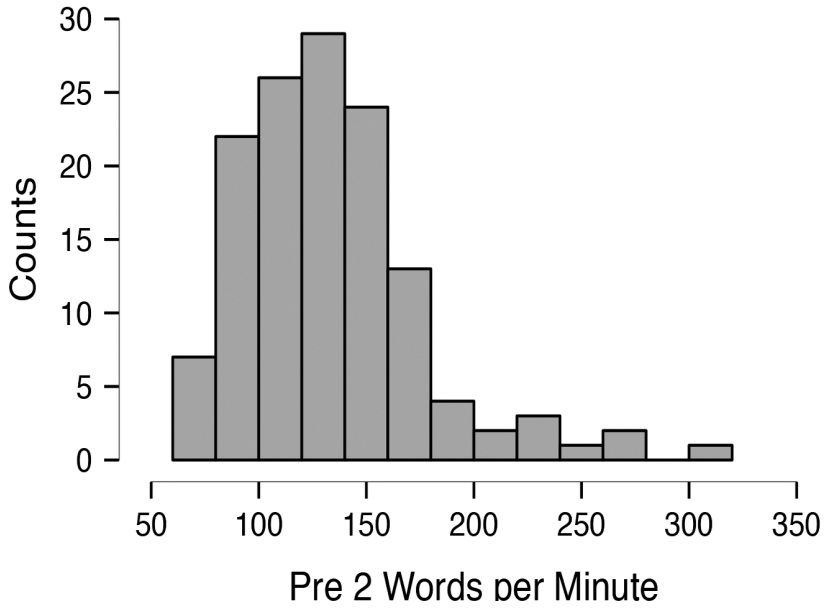

Figure 2. Distribution plots for Words Per Minute at Time 1 – Pre-test (n = 125).

RQ1. Did reading rate change from a pre-test to a post-test and then a delayed post-test?

For t-test (a), the Shapiro-Wilk test was significant suggesting that the pairwise differences were not normally distributed (W = .980, p = .001); therefore, the nonparametric Wilcoxon's Signed Rank Test was used. The test showed that there were significant differences between the Pre (2) Test WPM (Mdn = 125.8) and the Post (2) Test WPM (Mdn = 150.6) WPM (W = 592.5, p < .001). The rank-biserial correlation (r_B) = -0.846 suggests that this is a large effect size. See Figure 3.

For t-test (b), the Shapiro-Wilk test was significant suggesting that the pairwise differences were not normally distributed (W = .980, p = .048); therefore, the nonparametric Wilcoxon's Signed Rank Test was used. The test showed that there were no significant differences between the Pre (2) Test WPM (Mdn = 125.8) and the Delayed (2) Test WPM (Mdn = 125.1) WPM (W = 4442.5, p < .860). The rank-biserial correlation (r_B) = -0.018 suggests that this is a trivial effect size. In short, the pairs were different between the pretest and posttest, but not between the pretest and delayed posttest.

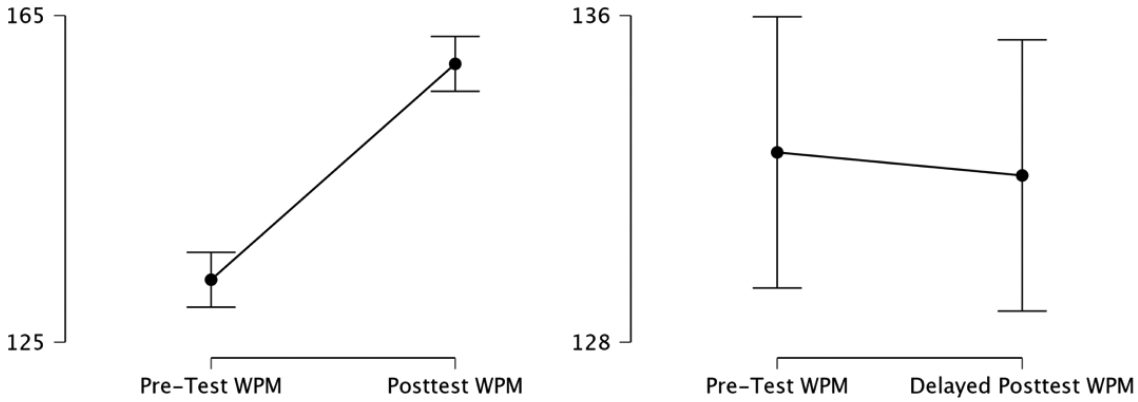

Figure 3. Words per minute (WPM) descriptive plots for paired-sample t-tests (n = 125).

RQ2. Does audio-assisted extensive reading lead to reading rate gains?

In a first run, two cases (i.e., students) were identified as having high standardized residuals (i.e., > 3.0). These two were temporarily removed and the data were rerun. In the second run, no additional cases were identified as problematic. The Durbin-Watson statistic, which checks for correlations between residuals, was 1.65. Ideally, this statistic should be between 1 and 3 (Goss-Sampson, 2022). The residuals vs predicted plot appeared to randomly distribute the residuals, suggesting that the assumption of homoscedasticity was not violated. The Q-Q plot of standardized

residuals showed that most of these residuals, except for the tails, fit along the diagonal. A simple linear regression was used, with bootstrapping (5000 estimates), to predict Delayed (2) WPM from Total Listening Time (TLT). TLT was shown to be nonsignificant: $F(1, 121) = 0.004$, $p = .949$, $R^2 = 0.000$. In short, delayed reading rate gains were not affected by AER. By contrast, a second linear regression with bootstrapping was run with total reading time (TRT) as the independent variable to predict Delayed (2) WPM. One additional case was identified as having high standardized residuals. This participant was temporarily deleted. TRT was shown to be significant: $F(1, 120) = 15.947$, $p = <.001$, $R^2 = 0.117$. In short, delayed reading rate gains were affected by total reading time but not total listening time.

In this current study, it was observed that reading rate improved between Time 1 (the pre-test) and Time 2 (the posttest), but not by the delayed posttest. That is, gains observed at the immediate end of the study did not carry over several weeks later. Previously, Gobel (2011) reported Japanese university student gains in TOEFL scores due to reading while listening over the course of one academic year. However, there are several weaknesses with that study. Over the course of one academic year, the participants in Gobel were not only enrolled in one listening course, in which they also did reading while listening activities, they were enrolled in three other required English-language courses for which the participants likely had many assignments that they completed over the course of the year. Thus, it is very likely that any gains in TOEFL scores, if at all, were not due solely to reading while listening. A second problem is that the reported TOEFL average gains were quite small. One might assume that many students did not improve beyond the standard error of difference and/or there was regression to the mean. Gobel did not report such data. Finally, the amount of reading while listening reported by Gobel was quite small, approximately four hours per semester, or 17,400 words on average, as Gobel noted: “this can hardly be called ‘extensive listening’” (p. 50). Thus, the claim that RWL improved TOEFL scores is weak, at best.

Milliner (2019) reported gains in TOEIC scores for participants who did reading while listening compared to a reading-only, and a control group. Importantly, similar to Gobel (2011), average gains were small for each of listening and reading for all three groups (from -15 to +22 scale points), which are within the unreported standard error of difference (±35 scale points). Importantly, there were large changes in standard deviations between groups and across time, potentially indicating the presence of outliers, but there was no report on how or if outliers were treated. Milliner also compared growth in listening vocabulary levels test (LVLT) scores for

these three groups but found no differences between groups. In the often poorly controlled settings of classroom-based research, it is difficult to claim gains in standardized test scores are the result of a particular research paradigm, in this case ER and AER. Instead, it might be more appropriate to focus on reading rate gains. Importantly, the current research was not interested in assessing the impact of reading while listening (or even extensive reading) on external standardized or other language tests, such as the LVLT; but whether audio assisted reading affected reading rate.

Among the participants in Chang and Millet (2015), those who read while listening increased their reading speed at a much higher rate compared to those who read only. Although the testing procedures in Chang and Millet and the current study were somewhat similar, the tests in Chang and Millet were approximately three times as long as those used in the current study. The longer texts likely led to greater accuracy in participants reading rate measurements, both within and between participants. Among other weaknesses in the current study, the relatively short tests is one weakness. These, and other weaknesses help to highlight the difficulties of measuring reading rate, and change in reading rate over time, in non-laboratory settings. Several of these weaknesses, and ideas for future research are described in greater detail below.

Limitations and Future Research

There are limitations of this study which lead to plans to improve the research. Firstly, reading rate measurements, pre, post and delayed, were problematic in several aspects. The six texts used for the reading rate measurements were chosen for similarity after evaluations of topic, word length and vocabulary difficulty. However, a post-facto analysis of the literary style found that the sixth reading, the delayed posttest reading, had a different style than the previous five readings. The sixth reading included questions and was not in the same narrative style as the previous five readings. While the participants' comprehension of this reading average was compared to the other readings, their reading rate was slowest. Possibly this new literary style affected reading rate. Another limitation with the reading rate measurements was that many students did not follow the read-write time-answer question practice which led to removing their data. The instructions were in English which may have limited participants' memory of the process, and the place to record reading time was not obvious. Secondly, a limitation was the control over students' reading. The average time spent listening to audio while reading was half of the time

spent reading. Ideally the reading and listening time would be similar. The time of the required audio-assisted reading was only 3.5 weeks which may limit being able to identify reading gains due to reading with listening.

Learning from the weaknesses of this study the writers plan to continue with the following improvements to increase reliability, control, and number of participants.

1. Improve reliability of reading rate measurement

As it seems that texts used in the delayed post-reading rate measurement were of unequal difficulty for students, we will take more care to choose texts. Texts will be compared not only by word-length and vocabulary level, but also by literary style. Following the selection of texts, we will pilot these texts on students not in the research. Also, the instructions on the front page of the reading rate measurement tests will be written in Japanese. Finally, the graphic design of the reading texts will be simplified to increase the completion rate of the reading rate measurements.

2. Schedule time for reading in class

Thirty minutes of class time will be given to students' reading fluency development. This time will ensure that a high percentage of students are consistently doing audio-assisted reading. It will also be used for reading rate measurements at the beginning and end of the quarter and explaining reading with listening benefits along with audio functions on the Xreading site.

3. Extend the time of audio-assisted extensive reading

Previously students were asked to include audio listening with their reading in the first half of the term (26 days). The time will be extended to the full quarter (52 days).

4. Increase the number of participants

The number of students participating in a future study could increase from eight classes which participated in this study to ten classes. This increase is expected to involve about 250 students.

Conclusion

This paper explained an attempt to measure the influence of audio-assisted reading on reading rate within an extensive reading program. In the middle of a year-long extensive reading program, students had an assignment to do audio-assisted reading. Reading-rate measurements were taken before and after the audio-assisted reading time period. Although there was an increase in reading rate it could not be attributed to audio-assisted reading treatment alone. Weaknesses in the method of measuring reading rate yielded inconclusive results, thus plans to improve

the research are suggested.

Acknowledgement

The writers would like to express gratitude to the reviewer who thoroughly evaluated the paper and gave many precise suggestions that improved the paper.

References

Beglar, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning*, 62(3), 665-703.

Campbell, A. P., Calman, R. & Campbell, J. (2021, August 13). Learner audio use in a large extensive reading program [Conference session]. ER Around the World Conference.

Chang, A.-S. (2011). The effect of reading while listening to audiobooks: Listening fluency and vocabulary gain. *Asian Journal of English Language Teaching*, 21, 43-64.

Chang, A. S., & Millet, S. (2015). Improving reading rates and comprehension through audio-assisted extensive reading for beginner learners. *System*, 52, 91-102.

Gobel, P. (2011). The effect of reading while listening on TOEFL gains. *Forum of Higher Education Research*, 1(1), 45-51.

Goss-Sampson, M. A. (2022). Statis*tical analysis in JASP: A guide for students, 5th Ed.* JASP. https://jasp-stats.org/wp-content/uploads/2022/04/Statistical-Analysis-in-JASP-A-Students-Guide-v16.pdf

Grabe, W. (2010). Fluency in reading – Thirty-five years later. *Reading in a Foreign Language*, 22(1), 71-83.

Huffman, J. (2014). Reading rate gains during a one-semester extensive reading course. *Reading in a Foreign Language*, 26(2), 17-33.

Kirchhoff, C., & Mision, M. (2022). Audio-assisted extensive reading: Learners' experience and attitudes. *The Reading Matrix*, 22(29), 1-11.

McLean, S., & Rouault, G. (2017). The effectiveness and efficiency of extensive reading at developing reading rates. *System*, 70, 92-106.

Milliner, B. (2019). Comparing extensive reading to extensive reading-while-listening on smartphones: Impacts on listening and reading performance for beginning students. *The Reading Matrix*, 19(1), 1-19.

Nation, P. & Malarcher, C. (2007). *Reading for Speed and Accuracy*. Compass Publishing.

Nation, I. S. P., & Waring, R. (2020). *Teaching Extensive Reading in Another*

Language. Routledge.

Park, A. Y. (2017). A comparison of the impact of extensive and intensive reading approaches on the Korean EFL learners' reading rate and reading comprehension development. *International Journal of Applied Linguistics & English Literature,* 6(3), 131-142.

Schmitt, N., Jian, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*. 95(1), 26-43.

Woodall, B. (2010). Simultaneous listening and reading in ESL: Helping second language learners read (and enjoy reading) more efficiently. *TESOL Journal*, 1(2), 186-205.

Appendix

Readings used to measure reading rate

Measurement Time	Reading Title	Standard Word Count
Pre 1	How We Learn	263
Pre 2	Step by Step Learning	261
Post 1	Reading Books	253
Post 2	The Library of the Future	261
Delayed Post 1	Fast Learners	282
Delayed Post 2	Different Ways to Learn	269