

One year of extensive reading on smartphones: A report

Brett Milliner

Tamagawa University
milliner@lit.tamagawa.ac.jp

Many research studies into Extensive reading (ER) have recognized its potential for improving language learners' reading proficiency and other language skills. However, some recent reviews of ER research (e.g., Waring & McLean 2015; Nakanishi, 2015; Mori, 2015) have argued for more refinement in research design so that ER research can be more effectively interpreted. Some areas of disconnection include (a) how extensive reading is conceptualized, (b) the lack of long-term studies, and (c) a lack of transparency concerning how much ER was done. Aiming to provide an evaluation of an ER treatment that responds to these three issues, this paper reports on 19 university students who read graded readers on their smartphones for the duration of one academic year (April 2015 to January 2016). Each student read an average of 263,767 words or 39 graded readers during this period. Changes in students' reading proficiency were evaluated in terms of changes in students' TOEIC® IP test scores and scores in the reading section of this test. Results showed that all but two students were able to reach the reading targets and the majority of the student sample were able to achieve an increase in TOEIC® scores after this treatment.

Keywords: e-book, extensive reading, MALL, smartphone reading

Introduction

The popularity of extensive reading (ER) components in tertiary English language programs throughout Japan is representative of the growing body of research advocating the benefits of this approach. Studies (e.g., Beglar, Hunt & Kite, 2012) have

empirically demonstrated the superiority of ER over other approaches – such as intensive reading – for developing second language students' reading skills. In his meta-analysis of 34 research studies in ER, Nakanishi (2015) concluded that ER “improves students' reading proficiency and should be a part of language learning curricula” (p.6). Another meta-analysis of note published by Jeon and Day (2015) focused on the impact of ER programs that addressed five core components: easy reading material; learners choose what they want to read; learners read as much as possible; reading is individual and silent; and teachers orient and guide their students (p. 302). Jeon and Day concluded that reading programs with this focus were able to achieve greater improvements in reading proficiency than intensive or traditional reading approaches.

Despite the many positive reports of ER in professional discourse, recent reviews of ER research (e.g., Mori, 2015; Nakanishi, 2015; Waring & McLean, 2015) have argued for refinement in research design so that results can be more effectively interpreted. Three key areas cited for refinement include: (a) how extensive reading is conceptualized (Waring & McLean, 2015) (b) the lack of longitudinal studies (Nakanishi, 2015), and (c) a lack of transparency concerning how much ER was done (Nakanishi, 2015; Waring & McLean, 2015). By presenting the case of 19 students reading graded readers as e-books on their smartphones for the duration of one academic year (April 2015–February 2016) this report evaluates the effects of ER in terms of the aforementioned guidelines for ER research.

Extensive reading

Most ER programs share a common purpose; that is, learners read large quantities of self-selected, simplified texts in an environment that promotes the enjoyment of reading in a foreign language (Day & Bamford, 1998; Renandya, 2007). There is, however, limited agreement among teachers and researchers concerning how ER is conceptualized (Waring & McLean, 2015) or the best way to implement it in the language classroom (Fenton-Smith, 2008). Also, there are some practical challenges facing teachers when implementing an ER component in their language class. These can include: the cost of creating and managing a graded reader library with titles covering a range of genres and reading levels (Day & Bamford, 1998; Hinkelman, 2013; Stewart, 2014), how to identify whether students are effectively reading (Brown, 2012; Campbell & Weatherford, 2013; Robb, 2015; Robb & Kano, 2013; Stewart, 2014), how to provide adequate reading support for students (Day & Bamford, 1998; Robb, 2009), and how to motivate students to read in large quantities (Brierley, 2009; Mori, 2015; Robb, 2002).

E-book extensive reading

An important distinction for the current study is that learners read e-books on their smartphones that were streamed from an online library. The researcher's motivation for this approach was based on an understanding that successful implementation of an ER component depends on students reading in vast quantities (Beglar & Hunt, 2014; Beglar, Hunt & Kite, 2012; Nation & Wang, 1999; Worden, 2015). Beglar and Hunt (2014) noted that students have to read 200,000 words or more per year for increases in reading fluency (e.g., development of sight vocabulary and processing of larger linguistic units) to be observed, while Nation and Wang (1999) recommend a much larger 500,000 word benchmark for vocabulary to be learned incidentally. As described by Mori (2015), the reality facing many

university language teachers in Japan is that most students do not like to read, even in their first language, and many live a very busy student life; hence most students are not so inclined to engage in ER. As a result, a teacher has to call upon a range of pedagogical techniques to get their students to read consistently (Robb, 2002). The researcher hoped that some of the advantages afforded by implementing an online e-book approach (listed below) would enable students to read vast quantities of text more consistently.

The lower costs of e-books, the possibility of sharing e-books through 24-hour online libraries, and convenient access from mobile devices represent an opportunity for teachers to more efficiently share reading resources with their students (Gerlich, Browning, & Westermann, 2011). Language learners' also stand to benefit from the affordances of e-books such as customized text display, glossaries, audio narration and dictionary access (Huang, 2013; Lai & Chang, 2011). Some e-book systems also generate analytics data, which provides readers with ongoing feedback on their reading (Brown, 2012; Huang, 2013). For example, a language learner could evaluate their: reading speed, reading time, reading level, total words read, and comprehension based on post-reading quiz results. What is more, teachers can utilize reading analytics data to identify students who might need extra reading support (Brown, 2012).

Xreading

For the facilitation of e-book reading, the program Xreading was adopted. Xreading is an online library of graded readers and learning management system (LMS) devoted specifically to ER. There are over 500 titles in the Xreading library, and an individual license (¥2,400 annually) grants access to the entire library, as well as the user analytics data and post-reading quizzes. The Xreading site is accessible via a mobile device or PC, and since it is a virtual library, multiple users or the entire class can read the same title at once. For teachers, administrators and students alike, the LMS generates an analytics' summary of reading progress, including words-per-minute (WPM) counts, total words read, books read, total reading time, and results of post-reading quizzes.

How ER was implemented in this study

In their individual meta-analysis of ER research, Nakanishi (2015) along with Waring and McLean (2015) recognized a wide degree of disconnect concerning each study's conceptualization of ER. To promote a more stable framework for research on ER, Waring and McLean (2015, p.165) presented a list of core and variable attributes of ER so that research in the field can be more effectively interpreted. In Table 1 below, the researcher adapted their ER core and variable dimensions (p.165) to illustrate how ER was implemented in this study.

Measuring ER effects using TOEIC scores

A number of researchers have used changes in TOEIC test scores to measure the effects of ER on Japanese university students' reading fluency and overall language skills. In their investigation into using ER as a means to promote greater reading speed, Storey, Gibson and Williamson (2006) compared TOEIC test reading section scores between a control group and a treatment group (N = 21) who did ER for eight weeks. Although the authors were able to observe an increase in student reading speeds, they were unable establish

Table 1. Summary of ER core and variable dimensions for this study (adapted from Waring & McLean, 2015)

Core Elements	Managed in this study by:
Fluent, sustained comprehension	Analysis of reading speed, post reading quizzes, ER awareness exercises (Brierley, Ruzicka, Sato & Wakasugi, 2010)
Large volume of material	Students were asked to read the following volume of words in each semester: <ul style="list-style-type: none"> ● Semester 1 Target 100,000 ● Semester 2 Target 150,000 To ensure that students read consistently, smaller monthly word targets were also used
Reading over extended periods of time	<ul style="list-style-type: none"> ● Students accessed graded readers via their personal smartphone or personal computer inside and outside the classroom ● Reading time recorded by Xreading system
Longer texts	● Students chose graded readers at a level appropriate for them
Variable Elements	
ER is conducted at home or in class	<ul style="list-style-type: none"> ● Students allowed 5–10 minutes of class time to access Xreading (60 classes in an academic year) ● Students instructed to read in their own time
ER is required	Students were required to read more than the semester word targets to receive 10% dedicated to ER in their final grade.
ER is monitored	<ul style="list-style-type: none"> ● ER was monitored by the Xreading system ● Student's WPM, level of book read and progressive word counts were scrutinized throughout the year
The reading is assessed	Students received a full 10% semester grade if they overcame the following benchmarks: <ul style="list-style-type: none"> ● Semester 1 Target – 100,000 words ● Semester 2 Target – 150,000 words
Reading follow-up activities	<ul style="list-style-type: none"> ● Post-reading comprehension quizzes (conducted inside Xreading) ● In-class discussions concerning books & reading progress
Graded or non-graded readers used	Graded-readers from Xreading library
Longer or shorter texts	Students were free to choose whichever text they liked in the Xreading library
Freedom to select texts	<ul style="list-style-type: none"> ● Students were free to choose whichever text they liked in the Xreading library ● No book could be read twice by the same student

a significant improvement in TOEIC scores for the treatment group. In their conclusion, Storey et al. indicated that their results may have been different if: (a) the treatment period was longer, (b) there was a larger number of participants, and (c) the amount and quality of reading done by students could be calculated more effectively. O'Neill (2012) investigated whether the effects of ER could be measured by TOEIC test reading section scores. Comparing the results between a treatment group who did ER (N = 213) and a control group

who did not ($N = 159$), no significant difference in test score gains was observed. Similar to Storey et al., O'Neill reported difficulty in establishing how much ER was done by participants, and whether the volume of reading done by the participants was actually enough for measureable gains to be observed. Sakurai (2012) looked at the relationship between the number of words read in an ER program and TOEIC test scores. In her study, 70 participants read an average of 46,531 words in an academic year. After comparing words read and post-treatment TOEIC test results, Sakurai concluded that the number of words read was significantly related to scores in the TOEIC test overall as well as scores in the reading section specifically. In her conclusion, however, Sakurai questioned the accuracy of the reading data, as students were required to maintain a reading log by themselves. The results observed in each of these three studies demonstrate some inconsistent findings relating to whether ER can influence gains in reading proficiency as measured by the TOEIC test. Moreover, each study was unable to reliably quantify how much reading students did. The researcher hoped that the reading data generated by the Xreading system would provide a more accurate measurement to evaluate ER's influence on reading fluency.

Methodology

Research questions

Participants in this study were asked to read graded readers using Xreading on their preferred mobile device for one academic year. As was noted in Table 1 above, students were required to read more than 100,000 words in the first semester and 150,000 words in the second semester to receive the maximum ER score of 10% for each semester's course. After one year of students reading graded readers on their smartphone or personal computer, this evaluation set out to answer the following research questions:

1. Did the implementation of the e-book system enable students to read in large volumes?
2. After reading 250,000 words, could students achieve an increase in overall and reading section scores for the TOEIC test?

Participants

This study was undertaken in the researcher's Four-skills English class at a private university in Tokyo, Japan. It was originally planned to track the reading progress of three classes over the duration of one academic year; however, the make-up of class members was changed after the first semester, leaving only 19 students. All participants were sophomore Japanese students aged between 19 and 21 years old studying in the Department of Humanities.

Measurement devices

Students read e-books on either their personal laptop computer, tablet or smartphone and reading analytics data was collected by the Xreading system. Overall results and results in the reading section of a TOEIC IP test taken before the course (December 2014) and at the end of the academic year (December 2015) were used to identify the effect of the treatment on students' reading proficiency. A t-test for dependent samples was conducted to compare these sets of scores. For the purpose of testing correlations between student's reading data

and TOEIC reading scores, Spearman's Rho correlation coefficient was adopted because of its ability to measure strength and association between ranked variables.

Results

Students' reading engagement

Table 2 (below) provides a summary of student engagement with ER over the duration of one academic year. As suggested by the average of total words read (263,767), almost the entire sample was able to reach the reading target, and some students continued to read thereafter. Only two students failed to reach the reading target by the end of the year. This result suggests that the implementation of mobile-based ER using Xreading combined with some pedagogical measures (e.g., grading rewards, in-class reading time, in-class ER discussions, and progressive word targets) enabled students in this learning context to reach the reading target. It should be noted here that students' Xreading logs revealed that for all participants, over 97% of their logins to the system were made from smartphones; hence one can assume that this sample almost entirely read on their personal smartphones.

Table 2. Summary of students' reading engagement for the 15-week semester:

Students	Average total words read per-student	Range	Standard deviation (words read)	Average reading time	Average number of books read (range)
n = 19	263,767	165,497–302,973	34,904	41 hours 25 minutes	39 (13–129)

Influence of reading volume on TOEIC reading scores

As part of the English class' course requirement, students sat a TOEIC IP test and the end of the academic year (December 2015). Comparing TOEIC test averages, this sample was able to achieve an average score 38 points higher after this ER treatment. A one-tailed t-test was able to confirm that this difference between scores was significant ($t = 3.18$, $p = 0.002$).

To evaluate whether students' ER engagement influenced their reading proficiency, students' reading section scores were compared with their reading scores achieved in December 2014. Students in this sample were able to achieve an average increase of 29 points in the reading section of the TOEIC test (see Table 4 below). Again, a significant difference was observed between these sets of scores in a dependent samples t-test ($t = 2.09$, $p = 0.027$).

Table 3. Summary of student performance in the TOEIC overall and TOEIC reading section:

	TOEIC December 2014	TOEIC December 2015	TOEIC Reading December 2014	TOEIC Reading December 2015
Student averages	454	492	183	212
& Range (n = 18)	175–650	315–650	60–305	85–350

Spearman’s Rho correlation coefficient was then computed to assess the relationship between the variables – changes in TOEIC scores and volume of words read (see Table 4). The calculation for the relationship between changes in student’s overall TOEIC and TOEIC reading score and the volume of words read resulted in no relationship being observed ($r = -0.17815$, $n = 19$, $p = .46559$). Similarly, no relationship was observed between reading volume and TOEIC reading score changes ($r = 0.07567$, $n = 19$, $p = .75817$). This result led the author to conclude that students’ ER may not have influenced the increase in students’ TOEIC scores.

Table 4. Summary of correlation statistics

Variable	1	2	3
1. Words read	-	-	
2. TOEIC score difference (December 2015–December 2014)	-.178		
3. TOEIC reading difference (December 2015–December 2014)	.076		-

Note. * $p < .05$. ** $p < .01$.

Discussion

Following 19 students reading graded reader books on their smartphones, this report attempted to address the three suggested refinements for research in ER. Following the guidelines set by Waring and McLean (2015), how ER was conceptualized was stated in Table 1. In this study students’ read for one year, a duration longer than a large majority of the studies analyzed in Nakanishi’s (2015) meta-analysis of ER research. The data generated by the Xreading system also allowed the researcher to establish a more transparent record of how much reading students did, especially compared to the student generated logs used in Storey et al. (2006) and Sakurai (2012). Moreover, the researcher was able to monitor post-reading quiz results and reading times to minimize the chances of student chicanery or cheating that was suggested in O’Neill (2012, p.32).

In response to the first research question, all but two students read over 250,000 words during the yearlong treatment. The students who failed to reach the target read 196,632 and 165,497 words respectively. The volume of reading done by this group is much larger than those reported in Sakurai’s (2012) and the other studies which investigated ER and its effects on TOEIC test scores. One explanation for students reading more are the conveniences an online library platform affords. In Cote & Milliner’s (2015) study investigating student perceptions of online extensive reading, students preferred online graded readers to paperbacks because they did not have to go to a physical library, and they could read anywhere anytime because they always carried their smartphones. Students in Cote and Milliner’s study also cited easier searching for books (i.e., they did not have to search through books on a shelf in a library) and the ability to track one’s reading progress as explanations for their preference for online ER.

Concerning the second research question it is difficult to commit to agreeing with the statement that after reading 250,000 words, students could achieve an increase in the TOEIC test scores and TOEIC reading section scores. This sample achieved a significant

increase in the TOEIC reading section and in the TOEIC test overall. These results differ from those observed by Storey et al. (2006) and O'Neill (2012) who failed to observe a significant change in scores after an ER treatment. However, this study was unable to establish a strong correlation between the volumes of words read and increases in test scores. Moreover, it is difficult to compare this study to ones in the past. In each of the aforementioned studies for Japanese learners, the ER treatment was different, students read significantly less, and the treatment period was shorter (e.g., Storey et al., 2006).

Some conditions prevent the researcher from concluding that this approach for ER is more or less effective than others. Firstly, this study looked at a sample of only 19 Japanese university students, which makes it very hard for generalizations to be made for other English students. A second limitation is that no comparison was made with a control group. After conducting his meta-analysis of ER research, Nakanishi (2015) advocated that researchers include a control group when a researcher is evaluating the effects of a treatment because researchers “need to have something to compare to” (p.29). Unfortunately, the breaking up of class groups after the first semester made the management of a control group impossible. In this case, a control group could have done no ER or read paperback graded readers. A final group of concerns relate to the researcher’s decision to use students’ TOEIC Test scores to evaluate reading fluency. While it is termed “reading test”, a variety of skills are tested in this section. For example, students’ knowledge of grammar patterns, vocabulary, test taking strategy and physical endurance are all evaluated as this 75-minute, 100-question reading test is preceded by a 45-minute, 100-question listening test. A better evaluation of students’ reading progress could involve: evaluating reading speeds before and after the treatment; evaluating changes in vocabulary size using the vocabulary size test (Beglar, 2010); or, changes in Extensive Reading Foundation Placement Test scores. Also, this study has not considered the influence of other TOEIC test study undertaken during the academic year. In the case of this sample, students completed drills and practice tests from a TOEIC test preparation textbook, and because TOEIC scores accounted for 20% of the overall grade for the class, individual TOEIC study may have had a stronger influence on the improvement of TOEIC scores.

Conclusion

This study reported on an English class’ extensive reading component whereby students read graded reader books on their personal smartphones for one academic year. At the end of this one-year pilot, all but two students read over the 250,000-word target. Results from a TOEIC test taken at the end of the treatment revealed most students achieved a significant increase in their overall and reading section scores. A more rigorous statistical analysis, however, was unable to support the thesis, reading a larger volume of words in ER would foster larger increases in TOEIC scores. Moving forward, the author intends to refine this research methodology and experiment with a new group of students.

References

- Beglar, D. (2010) A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101-118
- Beglar, D., & Hunt, A. (2014). Pleasure reading and reading rate gains. *Reading in a Foreign Language*, 26(1), 29-48.

- 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. doi:10.1111/j.1467-9922.2011.00651.x
- Brierley, M., Ruzicka, D., Sato, H., & Wakasugi, T. (2010). The measurement problem in Extensive Reading: Students' attitudes. In A. M. Stoke (Ed.), *JALT2009 Conference Proceedings*. Tokyo: JALT.
- Brown, D. (2012). Online support systems for extensive reading: Managing the tension between autonomy and institutional education. *The Language Teacher*, 36(2), 11–16.
- Campbell, J., & Weatherford, Y. (2013). Using M-Reader to motivate students to read. In S. Miles & M. Brierley (Eds.), *Extensive Reading World Congress Proceedings* (pp. 1–12). Seoul: Extensive Reading Foundation.
- Cote, T., & Milliner, B. (2015). Implementing and managing online extensive reading: Student performance and perceptions. *The IALLT Journal*, 45(1), 70–90.
- Day, R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. New York: Cambridge University Press.
- Extensive Reading Foundation (n.d.). ERF Placement Test. Retrieved from <https://erft.calps.shinshu-u.ac.jp/>
- Fenton-Smith, B. (2008). Accountability and variety in Extensive Reading. In K. Bradford Watts, T. Muller, & M. Swanson (Eds.), *JALT 2007 Conference Proceedings*. Tokyo: JALT.
- Gerlich, R. N., Browning, L., & Westermann, L. (2011). E-Readers on campus: Overcoming product adoption issues with a tech-savvy demographic. *Journal of Higher Education Theory and Practice*, 11(4), 41–52.
- Hinkelman, D. (2013). Blending technologies in extensive reading: MoodleReader in a Japanese university EFL program. In S. Miles & M. Brierley (Eds.), *Extensive Reading World Congress Proceedings* (pp. 91–100). Seoul: Extensive Reading Foundation.
- Huang, H. (2013). E-reading and e-discussion: EFL learners' perceptions of an e-book reading program. *Computer Assisted Language Learning*, 26(3), 258–281.
- Jeon, E. & Day, R. (2015). The effectiveness of core ER principles. *Reading in a Foreign Language*, 27(2), 302–307.
- Lai, J., & Chang, C. (2011). User attitudes toward dedicated e-book readers for reading: The effects of convenience, compatibility and media richness. *Online Information Review*, 35(4), 558–580.
- Mori, S. (2015). If you build it they will come: From a "Field of Dreams" to a more realistic view of extensive reading in an EFL context. *Reading in a Foreign Language*, 27(1), 129–135.
- Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, 49(1), 6–37.
- Nation, P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355–380.
- O'Neill, B. (2012). Investigating the effects of extensive reading on TOEIC® reading section scores. *Extensive Reading World Congress Proceedings* (pp. 30–33). Kyoto: Extensive Reading Foundation.
- Renandya, W. A. (2007). The power of extensive reading. *RELC Journal*, 38(2), 133–149.
- Robb, T. (2002). Extensive reading in an Asian context: An alternative view. *Reading in a Foreign Language*, 14(2). Retrieved from <http://nflrc.hawaii.edu/rfl/October2002/discussion/robb.html>

- Robb, T. (2009). The reader quiz module for extensive reading. In M. Thomas (Ed.), *Selected proceedings of the thirteenth annual JALT CALL SIG Conference 2008* (pp. 109–116). Tokyo: Japan Association for Language Teaching.
- Robb, T. & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language*, 25(2), 234–247.
- Sakurai, N. (2012). The relationship between the total number of words read and test scores in an extensive reading program. *ACTA Humanistica Et Scientifica Universitatis Sangio Kyotiensis Humanities Series*, 43, 112–122.
- Stewart, D. (2014). ER with online graded readers. *ERJ*, 7(2), 16–18.
- Storey, C., Gibson, K., & Williamson, R. (2006). Can extensive reading boost TOEIC scores? In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.), *JALT2005 Conference Proceedings* (pp. 1004–1018). Tokyo: JALT.
- Waring, R., & McLean, S. (2015). Exploration of core and variable dimensions of extensive reading research and pedagogy. *Reading in a Foreign Language*, 27(1), 160–167.
- Worden, D. (2015). A guide to setting appropriate reading targets for extensive reading. *The Center for ELF Journal*, 1(1), 55–60.

Author biodata

Brett Milliner is an assistant professor in the Center for English as a Lingua Franca at Tamagawa University in Tokyo, Japan. Brett's major research interests include mobile-assisted language learning (MALL), computer assisted language learning (CALL) and extensive reading/listening.