

Extensive Reading Using Authentic Materials with Elementary and Secondary EFL Students

Carol M Johnson, PhD

Renaissance, USA

Abstract

The benefits of Extensive Reading (ER) have been demonstrated by many research studies. Two outcomes are generally supported: 1) ER is effective for students beyond the beginning level of language proficiency, and 2) Graded Readers are critical to providing books that are comprehensible. The present studies ask the following questions: 1) Can ER be effective for students at the beginning level of language proficiency, and 2) Can authentic books be used in lieu of graded readers? Support for the present studies comes from Bahmani & Farvadin (2017) which reveals that students' comprehension improved when they read just above ($i+1$) and just below ($i-1$). Having enough books was critical to both studies. There needed to be enough books on enough topics at enough reading levels so that the participants not only *could* but *wanted* to read a lot. Participants also needed access to the books in and out of school, on and offline. As a result, authentic digital books were used. During the two studies, 2000 students in grades 2-9 averaged 10- and 42-minutes reading per week. Participants' baseline reading level was used to match them with books just above and below their reading level. Results from study 1 showed that 8 weeks with authentic digital books made up for nearly an entire year of English study in the classroom in reading, while the results from study 2 showed a one-half year gain in reading compared to native speakers of English. Details of the studies will be discussed.

Keywords: authentic books, CEFR A1-A2, comprehensible input, elementary EFL, reading

1 Introduction

The objective for students studying English as a second or foreign language is to understand, speak, read, and write in English. The number one challenge for teachers of English in a foreign language (EFL) setting is providing their students with adequate comprehensible input—listening and reading opportunities students can understand—to facilitate the acquisition of those four skills. In the case of reading, there are many choices. Educators need to know that materials they choose are both effective and efficient. In the end, they deserve to know that the books their students read facilitate real growth in English.

2 Literature Review

Providing EFL students with input that is both comprehensible and meaningful to the individual student matters. Input that meets both criteria is motivating, interesting, and leads to language acquisition. “We acquire language and develop literacy when we understand

messages, that is, when we understand what we hear and what we read,” when we receive ‘comprehensible input’ (Krashen, 2003).¹ Of the four skills, “reading is viewed as a crucial skill for foreign language learners to improve their language ability” (Chiang, 2015). Because developing high levels of language proficiency is dependent on a breadth and depth of vocabulary, reading has always been seen as a vehicle for providing students in a foreign language setting with the opportunity to engage with and build high levels of vocabulary necessary for use in future careers.

“Two major approaches have been used to develop reading skills: Intensive and Extensive Reading (Rashidi & Piran, 2011). Historically used in the classroom, characteristics of Intensive Reading include careful and close reading of a short text, with a focus on details of the text, including, but not limited to new vocabulary, grammatical forms, authors purpose, among others, etc. In the past several years Extensive Reading (ER) has “received special attention as an effective and promising way of developing foreign language abilities” (Yamashita 2013) as it “intends to develop good reading habits to build up knowledge of vocabulary and structure and to encourage a liking for reading” (Richards & Schmitt, 2010). Textbooks and lessons have limited impact due to exposure and language. The ability to provide large amounts of input is one reason ER is believed to affect language ability, as it is often the only source for preparing students to become independent readers; without ER it is difficult to acquire the academic vocabulary and depth of language they will need.

Common features of an Extensive Reading program include reading: 1) at the right level, 2) with comprehension, 3) a lot, 4) independently, and 5) silently. Thus far, much research been conducted about how best to implement an ER program, including at what level of language proficiency, how much vocabulary is learned, how much reading does it take required, the number of words, the number of books, as well as the type and of books, the level of books, etc.

Across the ER research, two conclusions have been drawn generally:

- 1) ER is effective for students beyond the absolute beginning level of language proficiency, and
- 2) Graded Readers are critical to providing books that are comprehensible.

The first acknowledges that students at the absolute beginning level of language ability do not yet know enough vocabulary to comprehend a lot of books independently. The second provides support for the notion that, for language to grow, the learner must engage with input just beyond his/her current language ability. Several studies, however, have shown that participants’ reading comprehension grew when students read just above ($i + 1$) and just below ($i - 1$) their language ability (Bahmani & Farvadin, 2017). The following two studies ask three research questions about Extensive Reading:

1. Do beginning EFL students demonstrate reading growth using authentic books when reading just above and just below their reading ability?
2. If yes, is the reading growth significant?
3. If yes, is the reading growth meaningful?

¹ Although comprehensible input refers to both listening and reading, the studies summarized in this paper focus on reading only.

3 Research Study 1

3.1 Method

Of the initial 2965, 902 beginning EFL students from 80 schools in grades 2-7 from Turkey participated in an 8-week study,² completing the pre-and posttest, and treatment.

The pretest measure was the students' scores on the Lexile Placement Exam, and the posttest was the last Lexile Assessment at the conclusion of the trial.³ The treatment was the exposure to the myON Reader™ books, and the posttest was the last Lexile Assessment score at the conclusion of the trial. myON Reader™ books, available in digital format, are authentic books, written with native speakers in mind, not for the purpose of teaching English.

Irrespective of actual grade, all participants were entered into the system as 1st graders to provide an appropriate entrance point to the Lexile baseline assessment for non-native English speakers. Participants then took the Lexile Placement Exam to determine their initial reading ability. Results of the assessment were used to match students with books that were 100 Lexiles below and 50 Lexiles above their reading level. Book recommendations were aligned to participants' age, reading ability, and interests. Time spent reading was determined by each classroom teacher.

3.2 Results

Research Question 1: Do beginning EFL students demonstrate reading growth using authentic books just above and just below their reading ability?

The results of the analyses of the Lexile scores showed that there were positive gains in the students' reading ability, on average, between the pretest and the posttest at each grade level. Averaging 42 minutes of reading per week, these gains increased as grade increased which is desirable. For example, with an average Lexile gain of 37, the gains in Lexile from Grades 2-7 ranged from 17 to 59, each grade's gains greater than the previous grade. Notably the greatest gains were from Grade 3 to Grade 4 (18 to 36) and from Grade 4 to Grade 5 (36 to 49).

Research Question 2: Is the difference between the first and last Lexile exam significant?

The gains were statistically significant overall and in all grades except in grade 3. The effect sizes, Cohen's *d*, indicate a small effect approaching a medium effect. As a general rule of thumb, *d*=.2 is a small effect, .5 is medium effect, and .8 is a large effect. In the case of this trial, the effect for grades 4-7 *approached medium*, while the effect for grade 2 was small. No effect size is presented for grade 3 as there was no statistical significance in the Lexile

² CEFR A1 and A2 are considered beginning levels of language proficiency. Students in the study were expected achieve CEFR A2 level by grade 7.

³ Designed by MetaMetrics, an independent third-party, for myON. Readability generally refers to how easy or difficult a text is to read and understand. Like most readability measures, sentence length and word difficulty are included in the calculation.

change. Overall, there were significant gains as a result of the trial as indicated by both statistical and practical significance. These results are quite promising given the number of students who took both the pretest and the posttest assessments *and* the short duration of the trial.

Research Question 3: Is the difference between the first and last Lexile exam meaningful?

Figure 1 provides more than simple pre-post, growth by grade level and average growth overall. It also depicts visually not simply overall growth during the trial, but the relationship between grades pre-and post, providing insight into how meaningful the difference was. For example, grade 3 post score (85) is higher than grade 4 pre score (66). This means that, in 8 weeks, grade 3 made up for nearly an entire year of English study in reading. The same is true for every grade-next grade in the trial. Grade 4 post (102) matches grade 5, grade 5 post (154) matches grade 6 pre (152), grade 6 post (208) surpasses grade 7 pre (181). (See Figure 1 below.)

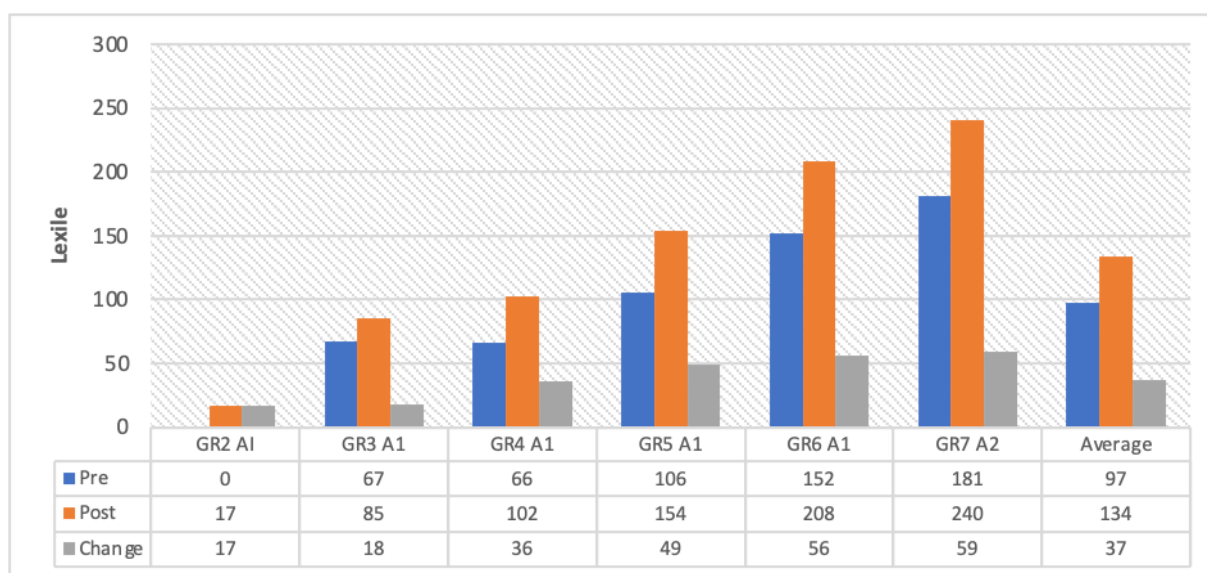


Figure 1. Average Lexile Growth (First to Last Week) by Grade Level

3.3 Discussion

The results of the analyses of the Lexile scores showed that there were positive gains in the students’ reading ability, on average, between the pretest and the posttest at each grade level. The overall average gain in Lexile score across grades 2-7 was 37, with gains that increased as the grade increased (See Figure 1). As a result, the response to all three research questions is “yes”, with one minor caveat for students in grades 2 and 3, who were in the very earliest of stages of language acquisition.⁴ Although both grades demonstrated growth, the gains were not statistically significant. This makes sense, as students must have some minimal level of any language before being expected to read it. However, a different pattern emerged from

⁴ Students in the study begin learning English in grade 2 with 2 hours of study per week, with 2 hours of English study in grades 2-4, 3 hours of English study per week in grades 5-6, and 4 hours of English study per week in grades 7-8.

grade 4 on. Grade 4 gain (36 Lexiles) was twice those of grades 2 or 3; gains for grades 5, 6, and 7 accelerated year by year (Lexiles 49-56-59, respectively).

4 Research Study 2

4.1 Method

1,104 beginning EFL students in grades 5-8 from Hungary participated in a 9-month study.⁵ completing the pre-and posttest, and treatment. Like Study 1, a pretest-posttest design was used to collect the data that were used to assess reading growth attained during the trial. The pretest measure was the students' scores on the Star Reading Assessment. Like other readability measures, the ATOS system used for Star Reading analyses sentence length and word difficulty. However, ATOS determines word difficulty by grade level, not word frequency.⁶ In addition, cognitive load--the type and amount of knowledge and thinking a student must have and use to answer test items correctly--is considered. As in Study 1, the treatment was the exposure to the myON Reader™ books, and the posttest was the last Star Reading Assessment score at the conclusion of the study.

All participants were entered into the system as 1st graders to provide an appropriate entrance point to the Star Reading assessment to provide a baseline for non-native English speakers.⁷ Participants then took the Star Reading Assessment to determine their initial reading ability. In addition to providing a reading level, Star Reading identifies a reading range--from just below to just above--for each student. The Zone of Proximal Development (ZPD) defines the readability range from which students should be selecting books to ensure sufficient comprehension and therefore achieve optimal growth in reading skills without experiencing frustration (p. 111 Star Reading Technical Manual). Teachers made book recommendations for students using ZPD guidelines. Three groups were formed during the research:

1. an **optimum** group, who used the program weekly and where the teacher possibly gave the students independent tasks between lessons
2. a **minimum** group, where the expectation was to use the program occasionally, at least once a month.
3. a **control** group, who was tested twice, but not use the program.

Midway through the study, it had already become clear that students were gaining from the use of myON. The decision was made to provide the control group access to myON, too. At the end of the 9 months, participants took another Star Reading Test.

⁵ Students began studying English in grade 4.

⁶ The ATOS system for evaluating the reading level of continuous text contains over 125,000 words in its graded vocabulary list. This readability formula was developed by Renaissance Learning, Inc., and designed by leading readability experts.

⁷ STAR Reading is a computer-adaptive test that determines a student's English reading level.

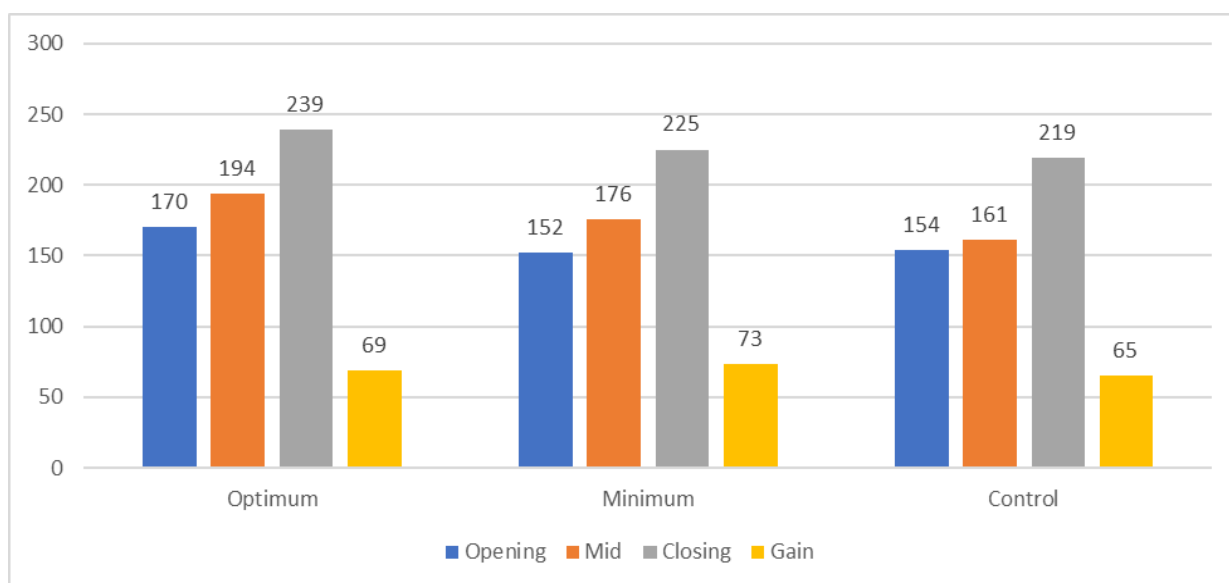


Figure 2. Average Star Scale Score Growth by Group

4.2 Results

Research Question 1: Do beginning EFL students demonstrate reading growth using authentic books just above and below their reading ability?

Averaging just 10 minutes per week for 9 months, there was a significant difference in the pre-posttests of the groups ($\chi^2(2) = 41.74, p = 0.000$), with mean rank values in the optimum group = 516.97, in the minimum group = 549.97 and in the control group = 402.5. However, the analysis of achievement results showed that there was no significant difference between the closing achievements of the groups, $\chi^2(2) = 5.87, p = 0.063$, with mean rank values in the optimum group = 486.19, in the minimum group = 473.21 and in the control group = 427.04. Figure 2 presents the English language scaled score grouped according to the measurement points. Average growth was 69 Star SS points.

Research Question 2: Is the difference between the first and last Star Reading Assessment significant?

Yes. The difference between the opening and closing Star Reading Assessment was significant ($p < .000$). Participants in the optimum group averaged a gain of 69 SS points; those in the minimum group gained an average of 73 SS points, while participants in the control group gained an average of 65 SS points. These gains were statistically significant for all three groups.

However, the difference in gains between the 3 groups did not meet that threshold ($p < 0.63$). There was no significant difference between gains of 69, 73, and 65 SS points, respectively.

Research Question 3: Is the difference between the first and last Star Reading Assessment meaningful?

Converting the Star SS to grade equivalent gives perspective on how meaningful the gains were. Each of the groups gained 0.4 grade using myON (See Figure 3).

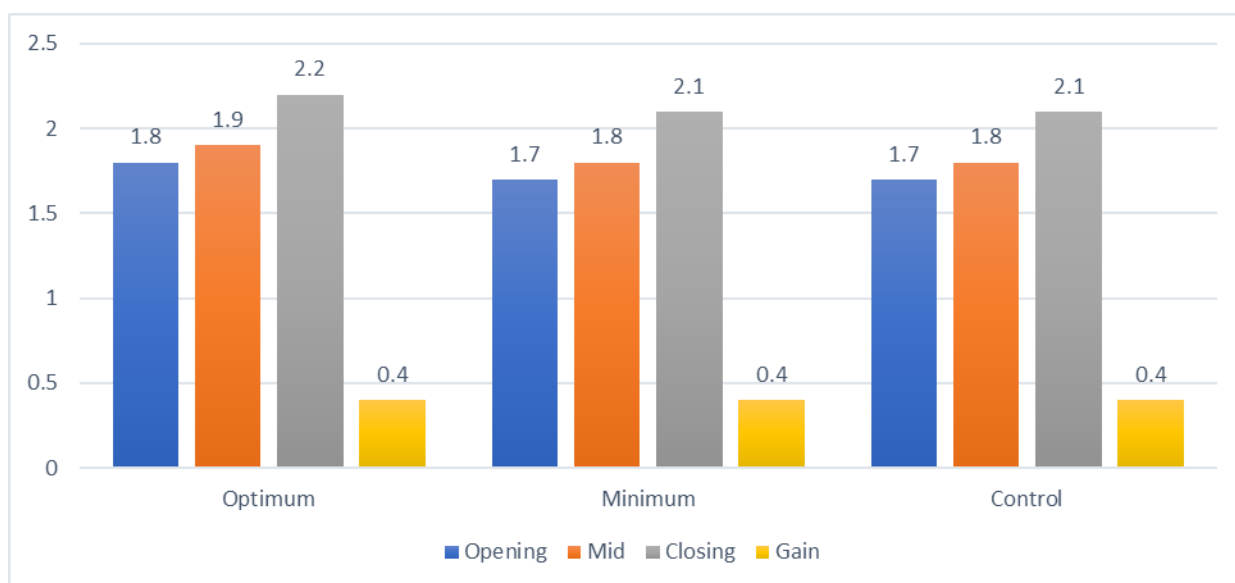


Figure 3. Average Gains by Group Converted to Grade Equivalency

4.3 Discussion

The results of the analyses of the Star Reading scores showed that there were positive gains in the students’ reading ability, on average, between the pretest and the posttest for all three groups, that the gains were significant, and they were meaningful. As a result, the response to all three research questions is “yes”.

5 Conclusion

Both studies demonstrate the efficacy of reading with authentic digital books for beginning students learning English in a foreign language setting.⁸ Reading with authentic digital books was not only effective, but efficient. In the first study students averaged 42 minutes per week over eight weeks, for a total of 5.6 hours of time spent reading. In that short amount of time, students gained one year of reading gains. In the second study, students averaged 10 minutes per week of reading over 9 months, for a total of 360 minutes or 6 total hours of time spent reading. In that short amount of time, students in all 3 groups gained nearly one-half grade equivalent as measured for native English speakers. Irrespective of the measure used, these studies demonstrate the critical importance of providing EFL students the opportunity to read authentic books that interest them. However, in the end, the ability to compare the reading level growth of students in the 2nd study to native English speakers provides a roadmap for a different outcome—a higher level of English proficiency—a level that may bring different opportunities in students’ futures.

⁸ However, based on the results of students in the first study in grades 2 & 3, while very early beginners demonstrate gains, significant gains occurred in grades 4-7 where students had already acquired more English.

References

- Bahmani, R., & Farvardin, M. T. (2017). Effects of different text difficulty levels on EFL learners' foreign language reading anxiety and reading comprehension. *Reading in a Foreign Language, 29*(2), 185-202.
- Chiang, M. (2015). Effects of varying text difficulty levels on second language (L2) reading attitudes and reading comprehension. *Journal of Research in Reading, 39*(4), 1-21.
- Krashen, S.D. (1985), *The Input Hypothesis: Issues and Implications*, New York: Longman.
- Krashen, S. (2003) *Explorations in Language Acquisition and Use: The Taipei Lectures*. Portsmouth, NH: Heinemann.
- Lennon, C., & Burdick, H. (2014). The Lexile Framework as an Approach for Reading Measurement and Success.
- Rashidi, N., & Piran, M. (2011). The Effect of Extensive and Intensive Reading on Iranian EFL Learners' Vocabulary Size and Depth. *Journal of Language Teaching and Research, 2*(2), 471-482.
- Richards, J. C., & Schmidt, R. W. (2002). *Longman dictionary of language teaching and applied linguistics*. Longman.
- Star Assessments for Reading Technical Manual*. (2021). Renaissance Learning.
- Schmitt, N., Jiang, X., & Grabe, W. P. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *Modern Language Journal, 95*(1), 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign language. *Reading in a Foreign Language, 25*(2), 248-263.