Evaluating an Extensive Reading Course

Kimberly Klassen and Tricia Allan
English for Liberal Arts
International Christian University

Abstract

An extensive reading (ER) component of an academic skills course for first year university Japanese L2 English readers is introduced. The rationale for ER is given: students can enhance their general reading skills, including reading speed and global comprehension, as well as improve word recognition automaticity. Course design draws on Nation's (2007) Four Strands for language development. Formative course evaluation is carried out to investigate: whether the main attributes of ER are met, and to what extent students benefit from ER. Data collection includes: word counts to determine amount of ER accomplished; timed reading scores to determine reading speed improvement; baseline vocabulary level testing; and student feedback. Preliminary data analysis indicates: a large amount of reading accomplished by participants; a significant increase in reading speeds; and a positive response to the course from students. The paper concludes with a few directions for future course evaluation.

The purpose of this paper is to introduce the Extensive Reading (ER) component of the Academic Skills course to faculty members and interested readers. The rationale for the course design is given, as well as preliminary data collection and findings pertaining to a formative evaluation of the ER component. A discussion follows focusing on the extent to which students might benefit from ER. The paper concludes with a look to future directions for both the ER component and its ongoing evaluation.

Rationale for an Extensive Reading Program

Extensive reading (ER) is defined by Grabe and Stoller (2002) as "reading in which learners read large quantities of material that are within their linguistic competence" (p. 259). To that end, graded readers are often used in ER programs because the vocabulary in these simplified books has been carefully controlled for. This allows a learner to choose appropriate reading materials to match their current proficiency level. Generally, it is agreed that 98% known vocabulary in a text is ideal for independent reading: this text coverage should allow for, but not necessarily ensure, adequate comprehension (i.e., 70% comprehension) (Nation, 2006; Schmitt,

Jiang & Grabe, 2011). Text coverage of 98% means that for every 100 running words, there are only two unknown words.

Day and Bamford (1998) suggest several features that characterize ER, including: participants select their own materials; reading is done for pleasure; and reading is its own reward (i.e., not assessed). However, some researchers have noted difficulty in meeting all the characteristics of ER as suggested by Day and Bamford (1998). For example, Hill (2013) offers several reasons why the aspect of pleasure should not be used to promote ER, most notably that many students find reading in their second language (L2) difficult and not enjoyable at all (p. 88). For that reason, Waring and McLean (2015) suggest four main attributes of ER to include: fluent comprehension, large amounts of texts, a high reading speed, and a focus on meaning. The authors suggest that other elements, such as whether reading is assessed, or whether it is pleasurable, could be considered variable elements of ER programs.

Discussing the importance of ER in a language curriculum, Grabe and Stoller (2011) argue that L2 readers "need to engage in reading for many hours at text- and task-levels appropriate to their abilities. It is only through extended exposure to meaningful print that texts can be processed efficiently and that students will develop as fluent readers (p. 24). Horst (2009) agrees, asserting that "key aspects of linguistic development cannot readily be acquired without [ER]" (p. 41). Much research has been carried out to investigate the potential benefits of ER. In his meta-analysis of 34 ER studies (N = 3,942), Nakanishi (2015) reports that experimental groups who participated in ER programs outperformed control groups in terms of reading proficiency with a medium effect size (d = 0.46). He also reports a large effect (d = 0.71) for pre-post contrasts, though correctly notes that because no control groups were used in these studies, results should be interpreted with caution. Other research has looked at how ER can improve learners' vocabulary knowledge (Nation, 2009, 2018), and enhance grammatical development (Day & Bamford, 1998). Furthermore, ER can improve learners' general reading skills, including their reading speed and overall comprehension (Nuttall, 2005). Perhaps an overlooked benefit is the potential of ER to promote automaticity (Hill, 2013), thus improving participants' lower-level processing skills.

Reading skills can be classified as involving either lower-level or higher-level processing. Lower-level skills are used to process text-based information and includes: word recognition and its sub-skills of phonological, semantic and orthographic processing; syntactic parsing; and building meaning from semantic and structural information (Grabe, 2009). In contrast, higher-level processing skills are reader-based, and concern the knowledge the reader brings to the text to build comprehension. For example, how the reader uses background knowledge, infers information and reacts to the text are all higher-level processing skills. Lower- and higher-level skills are thought to be interactive and hierarchical but not reciprocal: this means that for higher-level processing to be successful, lower-level processing needs to be automatic and efficient (Nassaji, 2014). Grabe (2009) notes that automaticity is important because word recognition and comprehension processes compete in the working memory: if too much time is taken to decode unknown vocabulary, then comprehension will begin to break down.

The development of learners' lower-level processing skills through ER should be considered vital in an L2 curriculum. As Perfetti (2007) explains, "In reading, the singular recurring cognitive activity is the identification of words. From this follow two other, related observations about reading: Comprehension depends on successful word reading. Skill differences in comprehension can arise from skill differences in word reading" (p. 357). Grabe and Stoller (2011) note the extensive time that is devoted to the development of lower-level skills in first language L1 classrooms. Because many L2 readers are already proficient readers in their L1, they are unlikely to need extensive training in higher-level processing skills. They are likely, however, to need many hours of L2 processing practice, especially when processing an L2 with a different writing system from their L1 (as in this context). By reading meaningful texts for extended periods of time, students will improve their word recognition skills necessary to become fluent, proficient L2 readers.

In order to maximize skills development in an ER program, a framework is used to guide the course design: Nation's (2007) Four Strands. A focus on comprehension forms the basis of the first three strands: *meaning-focused input, meaning-focused output*, and *fluency development*. In the fourth strand, *language-focused learning,* there is an explicit focus on language forms. *Meaning-focused input* is characterized by a low percentage of unknown items (i.e., matching the learner's current proficiency level) and a large quantity of material, following the 'time on task' principle of language acquisition (Nation, 2007). In an ER program, reading graded materials where 98% of vocabulary is known, provides the meaning-focused input. *Meaning-focused output* includes speaking and writing tasks that are demanding but achievable, with opportunities for practice. Example activities include pair/group speaking tasks, summarizing what was read. The third strand, *fluency development*, develops language skills to promote automaticity. The conditions should include familiar meaning-focused material, a pressure to perform faster, and regular opportunities for practice. An example activity is a timed-reading program in which: all the vocabulary is known; learners have a word per minute (WPM) goal to aim for; timed readings are done regularly; and learners record their reading speed progress over a period of time. The last strand is *language-focused learning*. Example activities could include maintaining a vocabulary record of new items, and practicing word cards in pairs. Each of the four strands is given an approximately equal amount of course and class time (i.e., 25%). Nation's (2007) framework draws on decades of research in second language acquisition for its rationale.

Course Design, Preliminary Data Collection and Analysis

In the academic year 2018, a new component of the Academic Skills (AS) course was implemented for Stream 4 students (CEFR A2/B1) in the English for Liberal Arts (ELA) program: Extensive Reading (ER). Before 2018, Stream 4 students read some short works of literature in their Reading and Writing (R + W) component of the AS course in the Spring term. A decision was taken in the ELA department to replace the R + W component with a component devoted to developing students' reading skills through extensive reading. A group of interested

faculty members (see Acknowledgements) met regularly in 2017 to articulate learning outcomes for the new component. The new ER component is required for all Stream 4 students in the Spring term. The Autumn and Winter terms are required for only those students who do not travel abroad for the Study English Abroad (SEA) summer program. In each term of 10 weeks, classes are held once a week for 70 minutes. There are approximately 20 students in each class in the Spring term, and between 15 and 17 students in the Autumn and Winter terms.

Nation's (2007) Four Strands were taken as a guiding principle for the course design for each term. Through this approach to course design, the four main attributes of ER as suggested by Waring and McLean (2015) are met: a large amount of reading, a focus on meaning, fluent comprehension, and a high reading speed. The Spring term focuses on introducing students to ER, including how to choose appropriate materials, and building reading speed. Instructors motivate the course for students by presenting goals and benefits of ER; instructors also model Sustained Silent Reading (SSR), and support engaging, active discussions. Approximately a quarter of class time is spent on SSR of graded readers selected by the participants (i.e., meaning-focused input). Another quarter of class time is given to reader diary entries, followed by oral summaries with a partner, of what was read that day (i.e., meaning-focused output). A quarter of class time is devoted to a timed-reading program (Quinn, Nation & Millett, 2007) (i.e., fluency development). Two timed readings are done in each class, at the beginning and end, as the students meet only once a week. Students record their scores, aiming for the fastest reading time with 70% comprehension. The goal is set to 300 WPM for the end of the term (Quinn et al., 2007). The final quarter of class time is spent on vocabulary cards and other word recognition activities (Crawford, 2005) (i.e., language development).

The Autumn term focuses on literary works of fiction, and continues to build students' lower-level processing skills. Students work in reading circles of five participants, reading and discussing the same graded reader. Discussions focus on elements of fiction, including themes. Each student is assigned a role: Discussion Leader, Summarizer, Connector, Passage Person, Word Master and Cultural Connector (for groups of six). Groups are responsible for selecting appropriate reading materials. The Winter term focuses on genres of nonfiction. Discussions focus on the importance of these topics to students as university students and citizens of the world. Less scaffolding is given in the Winter term for the discussions, and students are encouraged to do additional research on their topic.

In all terms, students demonstrate comprehension of what they have read by completing a short quiz on the website Mreader.org. The Mreader site is used provide SMART goals; that is, goals that are specific, measurable, attainable, results-oriented, and time-bound. Students need to get at least six out of ten questions correct on the Mreader quizzes. If they fail the quiz, they need to either write a summary of the book for the instructor, or read another book. As noted, 70% comprehension is considered ideal for independent reading. By setting the target to 60% for Mreader quizzes, we could account for any tricky or unusual questions. Participants' Mreader accounts keep track of word counts over the terms: the word count goal for the Spring term is set 70,000 words; Autumn, 100,000 words; and Winter, 130,000 words. This is the equivalent of one

or two graded readers at the A2/B1 level per week. Depending on their reading speed, this should take students about one hour per week (i.e., a reading speed of 200 WPM x 50 minutes = 10,000 words).

In order to evaluate how well the ER component is meeting stated objectives, data is being collected in four parts. First, word counts for each participant for each term in the academic year are recorded from Mreader.org to determine how much ER is accomplished. Second, to track participants' vocabulary growth over the academic year, baseline vocabulary levels were determined at the beginning of the academic year using the Vocabulary Levels Test (Schmitt, Schmitt & Clapham, 2001) at the 2,000 and 3,000 levels. The baseline scores will be compared to scores gathered in the last week of the academic year. Third, to track participants' reading speed development, word per minute (WPM) scores were recorded at Week 2 and Week 9 of the Spring term. Lastly, students' perceptions of the course were gathered through an online evaluation (Google Form) in the last week of the Spring term (see Table for evaluation questions). The evaluation was done in the last class to ensure a high response rate. Instructor feedback was not gathered: the two ER instructors were also course coordinators and authors of this paper. The evaluation consisted of closed and open ended questions, the latter of which were analysed using typological analysis.

Findings

As only the Spring term has passed at the time of the writing, the findings reported here represent a preliminary analysis.

Word Counts

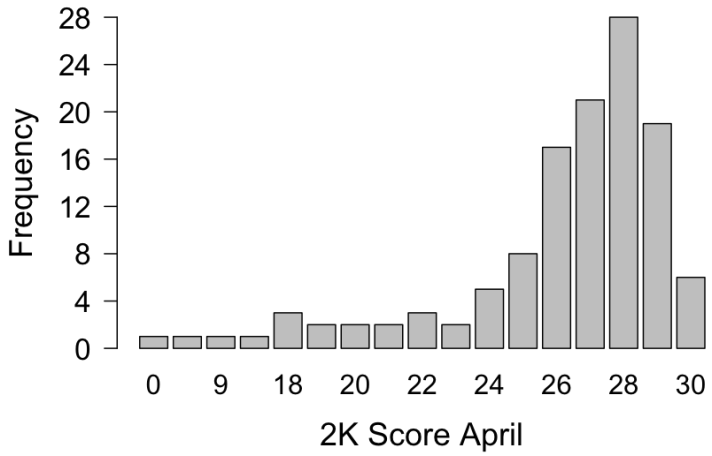
Word counts for each participant were recorded from Mreader.org to determine how much ER was accomplished (over each term and the academic year in total). Only partial results from the Spring term were available (n = 62). The word count goal for the Spring term was 70,000 words. Most students met this goal (M = 72,464, SD = 17,590), though there was variability in the word counts (Min. = 22,720, Max. = 140,956). Note that the top reader read double the given word count goal.

Vocabulary Levels

The Vocabulary Levels Test is used to track participants' vocabulary growth over the academic year. At the time of writing, only the Spring term baseline results were available (N = 122). These results will be compared to participants' scores at the end of the Winter term. Each band level test (2,000 and 3,000) has a possible 30 points. Scores of 27 or higher indicate that the student knows that level well; scores below 27 indicate that the level is not well known. As can be seen in Figure 1, many students had a good command of the 2,000 band (M = 25.88, SD =

4.49). However, there were 48 students who did not have a strong understanding of this band, which makes reading a difficult task for them, since the first 2,000 words will make up about 80% of the words in any given text.

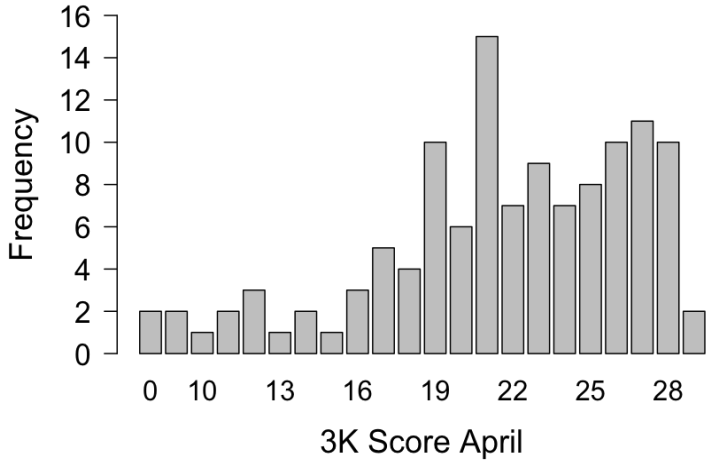
Figure 1. Frequency distribution of 2,000 level scores in April (*N* = 122).



Note. 2K = 2,000 word level

Figure 2 shows the frequency distribution for the 3,000 band scores (*N* = 121) (one student's data was missing). There was more variability in students' knowledge of this band level. Fewer students had a good understanding of the 3,000 band (*M* = 21.44, *SD* = 5.55).

Figure 2. Frequency distribution for the 3,000 band scores in April (N = 121).



Note. 3K = 3,000 word level.

Timed Reading Scores

Participants' timed reading scores (WPM) were recorded at Week 2 and Week 9 of the Spring term to determine their reading speed improvement. The scores were recorded where students had comprehension scores of 70% or better. A paired samples t-test (one-tailed) was run to compare scores from Week 2 to Week 9. The results indicated that scores were statistically significantly higher at Week 9 (M = 215.2 , SD = 55.06) than at Week 2 (M = 139.6, SD = 25.15), $t(104) = 15.986$, $p < .001$, with a very large effect size d = 1.56.

Course Evaluation

The main aims of ER in the first term were: building reading speed and introducing students to extensive reading, including how to choose appropriate materials. The Table shows students' feedback from the end of the Spring term using a Likert Scale for six closed questions.

Table

Results of Student Feedback on Course Evaluation Spring 2018 (N = 115)

Survey Questions	Likert Scale Responses (Percentages)				
	1	2	3	4	5
The course helped me to increase my reading speed.	1.7%	0.9%	7.8%	43.5%	46.1%
The course helped me to choose an appropriate book.	0	3.5%	18.3%	55.7%	22.6%
The course helped me to develop and build my vocabulary.	0.9%	5.2%	17.4%	52.2%	24.3%
The course helped me to gain world knowledge (e.g. new places, cultures, ideas) from the books I read.	0.9%	11.3%	24.3%	40.9%	22.6%
In general, there was a variety of interesting books to choose from in the library.	0	7.8%	20%	38.3%	33.9%
The Mreader website was easy to use.	2.6%	17.4%	19%	33.9%	27%

Note. 1 = strongly disagree, 5 = strongly agree.

Responses show a positive pattern. A large majority (89.6%) felt that ER helped to improve reading speed. Students also agreed that it helped them to choose an appropriate book; develop and build vocabulary; and gain world knowledge. In terms of resources and tools, most responded that there was a variety of interesting books to choose from and agreed that the MReader website was easy to use.

In terms of workload, students responded positively to the amount of reading both in and out of class with a large, majority noting that both were appropriate. Most (80%) replied that out of class reading was "appropriate" and 77.4% said that there was appropriate time given in class. Additionally, students were asked two open ended questions about the course. Analysis of student comments overall to "What points could be improved about the course?" is positive: 62% responded to that they did not want any change or with answers about which skills and attitudes were advanced by ER. Students who indicated a preferred change to ER also wanted more time for in-class reading. Ten (20%) desired more time for a specific, in-class task: five (10%) more vocabulary work; four (8%) more reading time and one (2%) more timed reading practice. An additional four students (8%) wanted more time for ER overall.

The remainder suggested changes in classroom management with seven (14%) requesting a range from "quieter" to "more teacher instruction" . A small number (3 r 6%) commented on changes to books available with some overlap with the 11 students (22%) who noted difficulties with the MReader website including when graded readers did not have a quiz. Some asked for specific content that is not a component of ER, for example:

(1) I think this course can more focus on toefl or Ielts basic vocabraly.(*sic*)

In response to "What were good points about the course?" the majority gave a positive comment with only one student replying "Nothing". Of the 114 (99%) who made positive comments, the largest number (51 or 44.7%) cited the amount of reading accomplished. A representative example is seen in excerpt 2:

(2) Thank(s) to the course, I read many books in English. As a result, my reading skill is improved.

Many students specified which reading skills had been improved. The greatest number (32 or 28%) noting reading speed in particular and a further 16 (14%) noting reading skills more generally, with an equal number citing vocabulary gains. After speed, the next most frequent comment (17 or 17%) was enhanced enjoyment of reading, with a range of answers noting changes in reading habits, being able to choose which books to read and gaining wider knowledge from their reading.

Discussion

As only the Spring term has passed at the time of writing, the discussion necessarily focuses on that term. The ER component seems to have met the four core attributes of ER as outlined by Waring and McLean (2015): a large amount of reading, a focus on meaning, fluent comprehension, and a high reading speed. Regarding the first, a large amount of reading was done by the majority of students in the Spring term; most students met the goal of 70,000 words. Several students reached this target by Week 6 of the term, which prompted an increase to the Autumn and Winter goals. Higher targets should motivate faster readers while still being attainable for slower readers. As noted above, a student with a reading speed of 200 WPM could meet this goal by reading 50 minutes per week. Also, the inclusion of 20 minutes of SSR in class in the Spring term should help slower readers to meet the target.

Another central attribute of ER is a focus on meaning. This aspect was met not only through the Mreader quizzes, which focuses on global comprehension, but also from the classroom activities of reader diaries and sharing in pairs, summarizing what was read. The core feature of fluent comprehension was also met by using Mreader: as noted in the course design, students need to get at least 60% correct on the Mreader quizzes.

The last attribute of ER is a high reading speed. This feature was ensured through the timed reading program. By the end of Spring term, the average reading speed was 215 WPM. This was statistically significant higher than the average of 140 WPM at Week 2. This finding suggests that participants' reading speed improved greatly as a result of the timed reading training, even over a short period of eight weeks.Thus, we can be confident that the word count goals of 100,000 words in the Autumn and Winter terms will be met. Also, 215 WPM is a very good reading speed for ELA students to attain. In comparison, most L1 English readers read at speed of 300 WPM (Nuttall, 2005). Therefore, Stream 4 students will be well prepared for the reading components of future courses in the College of Liberal Arts (CLA).

Regarding students' perceptions of the ER component, students consistently identified that ER classes helped them to develop reading skills and fluencies and clearly saw a positive cost benefit analysis of doing ER. Findings support student awareness of the main aims and positive outcomes for their own learning and skills. Out of class work was designed to require approximately an hour. One third scored their own, weekly, averages at 30-60 minutes and at 61-90 minutes spent on ER. Yet, overall students felt the time required was satisfactory.

Findings show that students identified which skills were improved by ER, wanted more time on specific skills and some went further, suggesting more time for ER overall. As seen in excerpt (3) they commented on developing new, positive attitudes to reading. Excerpt (4) illustrates being positive about the opportunity to read widely and its benefits. Quotation (5) demonstrates that ER is seen as a counter balance to difficulties with university studies more widely.

(3) In April, I felt resistance about reading English books. But, I do not feel resistance

about English books and I like English books thanks to ER class.

(4) Before coming ICU, I have few opportunity to read English books. So thanks to this environment, I get used to reading English books. Also, I could improve my speed reading skill.

(5) This university is very hard everyday, so this course is my refresh times. Because this course gave me time to read books. In addition, I was able to read a book of a new genre.

Excerpt 6, shows that students understood why teachers advised how to choose books at a particular level for ER, despite some initial reluctance to choose books that were seen as "easy":

(6) It was good I found my favorite books in ICU library. Choosing an appropriate book has a big effect on how hard we work on reading.

Aspects of class management, resources and tools as noted by students have been considered and changes made. Most are minor, such as revising information and advice for teachers to troubleshoot in advance of issues arising. Library staff have been supportive of changes to shelving to showcase the most appropriate books for ER whilst maintaining access, online to those less suitable. Additional graded readers have been purchased. The different foci for autumn and winter terms will give those students who want to read different genres the opportunity to do so.

Overall, student responses in feedback demonstrate an understanding of their own learning and how ER in and beyond classes has supported them to develop skills and acquire knowledge. Excerpt (7) is indicative of a number of responses regarding reading speed. It implies satisfaction with learning gains and a sense of ownership to continue making further progress.

(7) I improved my speed reading skills. First I could not read quickly. However I can read very quickly. Because my Time Reading score is 367! So I would like to improve my reading speed more and more.

It is particularly positive that students recognized the ways in which their reading had improved and some noted their intentions to continue reading widely, beyond the class.

Conclusion

Considering the extent to which students might benefit from ER, this preliminary analysis indicates a few positive outcomes. First, a statistically significant increase was seen in reading

speed among the participants. This improvement should prove beneficial in their other academic courses. Second, students' feedback on the ER component indicated the majority recognized these benefits for their reading skills, also noting wider gains in increased motivation and opportunity to read with some making the connection that transferable skills acquired support learning more broadly.

It is predicted that students' vocabulary will improve as a result of ER; however, the data will need to be analysed at the end of the academic year. Of course, an improvement in vocabulary knowledge cannot be attributed to only the ER component, as students take part in many other ELA courses. However, we do hope to compare the vocabulary scores of students who participate in the ER component all year with those students who participated only in the Spring term to determine if there is any effect on vocabulary knowledge.

Limitations to the preliminary analysis presented here include the nature of the feedback given in the course evaluation. While the feedback from students was overwhelming positive, it should be noted that the course evaluation was conducted in the classroom with the instructor present to ensure a high response rate. However, the presence of the instructor may have skewed comments toward the positive.

Finally, the authors of the paper hope that the ongoing evaluation of the ER component will add to the body of literature on ER. They seek comments from interested faculty and readers on the course design and evaluation.

Acknowledgements

Jennifer Yphantides, as coordinator of ARW Stream 4, set up the ER component group who met during academic year 2017 to discuss and plan the program: Tanya Erdelyi, Simon Evans, Akiko Fukao, Chris Gallagher, Joe Garner, Ged O'Connell, and the authors of this paper. Mari Aitoku from the ICU Library provided much support; requests for graded readers appropriate for Stream 4 students were fulfilled. Thanks also to the students who gave their feedback on the component and shared their timed reading scores.

References

Crawford, M. (2005). Adding variety to word recognition exercises. In *English Teaching Forum* (Vol. 43, No. 2, pp. 36-41).

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Grabe, W. (2009). Reading in a second language: Moving from theory to practice. Cambridge: Cambridge University Press.

Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. Harlow, England: Pearson Education.

Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). England: Pearson.

Hill, D. R. (2013). Survey review: Graded readers. *ELT Journal, 67*(1), 85-125. doi:10.1093/elt/ccs067

Horst, M. (2009). Developing definitional vocabulary knowledge and lexical access speed through extensive reading. In Z. Han & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 40-63). Michigan: University of Michigan.

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly, 49*(1), 6-37. doi:10.1002/tesq.157

Nassaji, H. (2014). The role and importance of lower-level processes in second language reading. *Language Teaching, 47*(01), 1-37. doi:10.1017/s0261444813000396

Nation, I. S. P. (2006). How large a vocabulary size is needed for reading and listening? *The Canadian Modern Language Review, 63*, 59-82. doi:10.3138/cmlr.63.1.59

Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1*, 2-13. doi:10.2167/illt039.0

Nation, I. S. P. (2009). *Teaching ESL/EFL Reading and Writing*. New York: Routledge.

Nation, I. S. P. (2018). Reading a whole book to learn vocabulary. *ITL - International Journal of Applied Linguistics, 169*(1), 30-43. doi:10.1075/itl.00005.nat

Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford: Macmillan Education.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi:10.1080/10888430701530730

Quinn, E., Nation, I. S. P., & Millett, S. (2007). *Asian and Pacific Speed Readings for ESL Learners: Twenty passages written at the one thousand word level*: English Language Institute Occasional Publication No. 24.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*, 26-43. doi:10.1111/j.1540-4781.2011.01146.x

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55-88. doi:10.1177/026553220101800103

Waring, R., & McLean, S. (2015). Exploration of the core and variable dimensions of extensive reading research and pedagogy. *Reading in a Foreign Language, 27*(1), 160-167.