Reading in a Foreign Language April 2011, Volume 23, No. 1
ISSN 1539-0578 pp. 1–16

Extensive reading and the effect of shadowing

Takayuki Nakanishi
Tokiwa University
Japan

Atsuko Ueda
Ibaraki University
Japan

Abstract

The aim of this study is to investigate the effects of extensive reading (ER) and shadowing on performance on reading comprehension tests. This study addressed the following research questions: (a) Can extensive reading improve students' reading comprehension? and (b) can shadowing enhance the effects of extensive reading? The participants in the study were 89 Japanese university students majoring in human science. Based on two experimental groups and two control groups, we examined the relationships and interactions of the two variables (ER and shadowing) over a one-year treatment (two semesters), using ANOVA. Three reading comprehension tests, a pretest, posttest 1 (after the first semester), and posttest 2 (after the one-year treatment), were administered. The results indicated that there was no statistically significant difference among groups, but a significant difference was found between the three test scores. Results are also considered in terms of an increased understanding of shadowing, and implications for curricula and classroom applications are discussed.

Keywords: ANOVA, extensive reading, reading comprehension, shadowing, SLEP

Extensive reading (ER) has been gaining popularity in English Language Teaching (ELT) settings in Japan. In ER classrooms, students read a "huge amount of very simple text so that [they] can read smoothly, confidently and pleasurably" (Waring & Takahashi, 2000, p. 6). Nuttall (2005) described ER as "the easiest and most effective way to improve [students'] skills" (p. 127) and claimed that it is "much easier to teach people to read better if they are learning in a favorable climate" (p. 127).

Many English as a foreign language (EFL) researchers (e.g., Camiciottoli, 2001; Day & Bamford, 1998; Krashen, 1982; Mason & Krashen, 1997; Robb & Susser, 1989; Yamashita, 2004) have suggested ER as a good strategy to improve reading proficiency, and a large number of studies (e.g., Elley & Mangubhai, 1981; Greenberg, Rodrigo, Berry, Brinck & Joseph, 2006;

Hafiz & Tudor, 1990; Lai, 1993) have confirmed its effectiveness in building linguistic competence (e.g., reading ability, vocabulary, writing and spelling skills). Numerous classroom activities using graded readers have also been described (e.g., Bamford & Day, 2004; Nakanishi, 2005).

Throughout the world, research has been conducted on ER and many studies have focused on the development of reading fluency. Nation (1997) stated that reading simple stories under appropriate time pressure is effective in helping students to gain reading fluency. The latest study by Iwahori (2008) examined the effectiveness of ER on reading rate and cloze test scores. Thirty-three Japanese high school students were provided with graded readers as homework for seven weeks. The results indicated that ER improved reading fluency and general language proficiency.

Researchers have also investigated how an extensive reading program can be utilized in a classroom for adults who have difficulty reading (Greenberg et al., 2006). In this study 22 out of 27 participants were native English speakers. They showed gains in reading fluency and expressive vocabulary, whereas no gain was found in their receptive skills. Nonetheless, many of the participants enjoyed the program and expressed a joy of reading. The lack of control groups in the above-mentioned studies, however, makes it difficult to attribute these effects to ER because intervening factors (e.g., exposure to reading outside of class and particular classroom experiences) may have played a role.

Other studies (e.g., Horst, 2005; Lai, 1993; Takase, 2007a) also investigated the effects of ER without control groups. The tendency of ER researchers to conduct studies without control groups cannot be overlooked. Two more recent studies on ER (Kweon & Kim, 2008; Yamashita, 2008) are other representative cases. Kweon and Kim (2008) were interested in incidental vocabulary acquisition and retention in ER. In their study, 12 Korean learners underwent three vocabulary tests. The results showed a statistically significant difference between the pretest and posttest 1. Furthermore, the words learned were retained one month later. To explore the effects of ER on different aspects of second language (L2) ability, Yamashita (2008) conducted a study in which 31 Japanese university students underwent a 15-week ER course. The results indicated that the strength of ER tends to be revealed in terms of general reading ability, and that linguistic abilities, such as vocabulary, spelling and morphosyntax, may appear at some point in the future.

Furthermore, ER can play another important role in the classroom. For example, in ER programs, students are highly motivated and may develop a positive attitude toward reading in the L2 (Takase, 2003; Ueda, 2005). Nuttall (2005) depicts the cycle of frustration and the cycle of growth (see Figures. 1 and 2)[1] and insists on encouraging students to enter into the virtuous circle of the good reader. As Figure 2 suggests, once a reader understands stories, he or she will enjoy reading, which results in increased reading frequency and, ultimately, reading more books. If a teacher provides easy-to-understand books, then even a reluctant student can enter the virtuous circle of the good reader in her model.

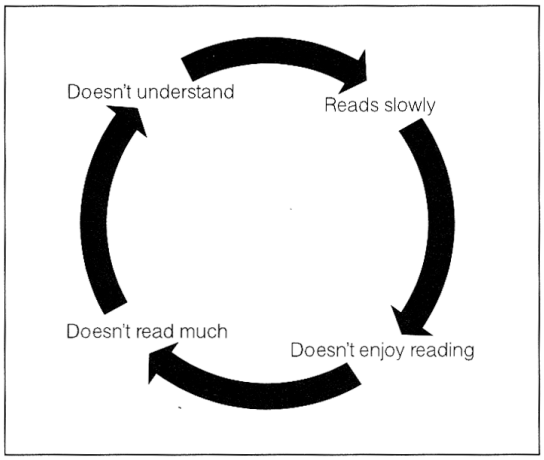

Figure 1. The vicious circle of the weak reader

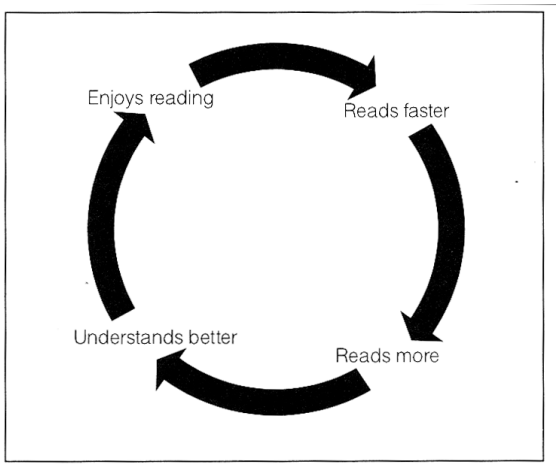

Figure 2. The virtuous circle of the good reader

In terms of implementing ER, Sakai and Kanda (2005) delineated three golden rules for successful ER teaching: (a) Students should not use a dictionary; (b) if students encounter some unfamiliar words, they may simply skip them; and (c) students can quit in the middle of reading if they find the book uninteresting and can switch to another book. Sakai and Kanda (2005) and Takase (2004) also strongly emphasized the importance of implementing an ER program not just outside the classroom, but inside as well.

Starting with very simple books, whatever the students' level may be, is another crucial element of an ER program. It may be controversial to allow university students to read books intended for very young children because the academic level might not appear suitable. However, most students do not remain at the lowest levels for very long; they improve their reading skills and in a short period of time move to higher levels of books. There are two major reasons to have students read very easy books: (a) An ER program promotes students' reading fluency, so it is crucial to read books that are much easier than the materials they use in other English programs; (b) due to the differences between the reading systems of English and Japanese and the translating habits developed in L2 education, most students are unable to read English without translating into Japanese word by word. Reading simple English at a fast pace makes it easier for students to translate less and possibly cease translating from English to Japanese in their mind.

In regular ER class, the students were responsible for selecting and reading books. The teacher walked around the classroom, helping students choose books, consulting, advising students on reading or having short discussions about their books, sharing and exchanging impressions and sentiments—all on an individual basis. When students approached the displayed books to choose new ones, the teacher was there and talked a little about each of the books they examined. It was very important that the teacher knew something about (or, preferably, had read) most or all of the books in order to give appropriate, funny, and motivating comments to each student. This is a very important role for the teacher in an ER class. One of the concerns among those language teachers who hesitate to employ ER is the "difficulty of the different role of the teacher" (Takase, 2007b, p. 8), or, as Muto (2006) asked, "What do teachers do? They don't appear to be

teaching" (p. 11). By way of explanation and in response to this reluctance, it is important to note that in ER classrooms, teachers do not "teach" per se but rather facilitate the students individually and establish and maintain the conditions of a reading environment that encourages ER.

Shadowing

Shadowing was initially developed as a way for training simultaneous interpreters, but currently many junior high school and high school teachers are adapting the techniques to their language classrooms. Research on shadowing has begun to appear in journals (e.g., Nye & Fowler, 2003; Ota, 2007); however, shadowing is still an under-investigated research area in applied linguistics. Shadowing is defined as an act or task of listening in which the learner tracks the target speech and repeats it immediately as exactly as possible without looking at a text (Kadota & Tamai, 2004). Over an extended period of time, students perform shadowing with various materials, which can affect brain processing. Déjean Le Féal (1997) stated, "shadowing is a good way to improve a foreign language precisely in that it draws attention to every single word of an utterance, especially structure words which normally do not even register when heard" (p. 621). It also provides students with sufficient input aurally.

Kadota (2007) suggested shadowing as a good way to reproduce English prosody. From a cognitive psychological point of view, Kadota illustrated how shadowing could automatize speech perception and also internalize new items. He thoroughly distinguished the difference between reading aloud and shadowing, not just in training methods, but in terms of efficacy—aural reading promotes the automatization of written lexical access, not speech perception.

To sum up, identifying the effects of ER requires control groups for comparison of results. Many ER studies have lacked control groups, which made it harder for us to determine whether the claimed effects resulted from the ER treatment or not. More research on shadowing should be conducted in order to recognize what constitutes shadowing and its effects. Thus, this study attempts to investigate the effect of ER and the interaction effect between ER and shadowing, using two control groups. Therefore, the following two research questions were addressed:

Research Questions

1. Can extensive reading improve students' reading comprehension, as compared to control groups?
2. Can shadowing enhance the effects of extensive reading?

Method

This section describes the books we used when conducting ER, the ways in which shadowing was implemented, and the instructions given to students in each class.

Participants

The study was conducted with 89 first-year Japanese university students aged 18-20 years old who were majoring in human science. Four intact classes—two experimental and two control groups—were compared. Twenty students attended an ER-only class (Group 1) and 22 students attended an ER-and-shadowing class (Group 2). The two control groups (Groups 3 and 4) were translation-based classes (n = 21 and n = 24, respectively). These two control groups were taught using a traditional translation method in which students were given a short English passage to translate into Japanese and asked to answer comprehension questions concerning the passage. In the translation-based class, the students were required to translate about two to three paragraphs into Japanese every week. They were not required to read books outside of the class or write book reports. The students in the control groups read approximately four to five pages a month.

In addition, all four groups were taking listening-based classes conducted in computer-assisted language learning (CALL) classrooms in which a teacher presented a variety of listening activities, such as answering multiple choice questions and dictations, but no reading assignment. All the groups attended 30 class sessions each lasting for 80 minutes. They all took a reading comprehension test three times. Initially there were 100 participants for this study. Individuals who missed one or more of the tests or who did not answer more than half of the items were excluded. After Test 2, the number decreased to 93. After Test 3 at the end of the academic year, another four participants had to be eliminated. Thus full analyses were performed on 89 participants.

Instruments

Reading comprehension test. The Secondary Level English Proficiency Test (SLEP), developed by Educational Testing Service (ETS, 2003), was used to check participants' L2 comprehension. SLEP was chosen following two previous studies (Takase, 2003, 2007b). The test contains listening and reading sections, and there are three equivalent types of the test—test forms 4, 5 and 6. In this study, only the reading parts of the three test forms were administered. To determine the statistical characteristics of the test forms and to equate the forms to the current SLEP scale, these three types of tests were originally piloted on 1,650 nonnative English-speaking students. The reliabilities of the reading sections were estimated using Cronbach's coefficient alpha as ranging from .88 to .91. The reading part has 71 items, and the SLEP scale ranges from 10 to 35; however, following the same line of scoring by Takase (2003), this study used the raw scores.[2]

The test takes 45 minutes to complete. All questions are multiple-choice. For items 1-12, students are given 12 sentences and a picture of four people and are expected to match the right person to the right sentence. For example, if a girl is thinking about becoming a pianist, then the student should choose the sentence *I want to be a pianist*. For items 13-28, each item shows four different pictures and a sentence. The students are required to choose the picture that best illustrates the sentence. Items 29-35 are grammatical questions. Within a sentence, a word is missing, and four possible choices are given. The students are asked to choose the one that best completes the sentence. For items 36-40, comprehension questions are provided based on a passage. Students are expected to choose the best answer. Items 41-45 and 52-59 are the same types as items 29-35, while items 46-51 and 60-63 are the same as items 36-40. For items 64-71,

long passages are provided, and students have to answer several comprehension questions for each.

Procedures

At the beginning of the course, test form 4 (Test 1) was administered to all four groups, and after four months of instruction, test form 6 (Test 2) was administered. Then, at the end of one full academic year of instruction, test form 5 (Test 3) was administered. The course descriptions are displayed in Table 1. Prior to analysis, assumptions of normality, homogeneity, and linearity were examined following recommendations found in Green and Salkind (2005).

Table 1. *Course Descriptions (30 classes over one year of instruction)*

	Experimental Group 1	Experimental Group 2	Control Groups 1 & 2
Instruction	ER	ER & Shadowing	Translation
Class 1	Test 1 (Pretest) SLEP (Form 4)	Test 1 (Pretest) SLEP (Form 4)	Test 1 (Pretest) SLEP (Form 4)
Classes 3-13	ER Instruction	ER & Shadowing Instruction	Translation-based Instruction
Class 15	Test 2 SLEP (Form 6)	Test 2 SLEP (Form 6)	Test 2 SLEP (Form 6)
Classes 16-29	ER Instruction	ER & Shadowing Instruction	Translation-based Instruction
Class 30	Test 3 SLEP (Form 5)	Test 3 SLEP (Form 5)	Test 3 SLEP (Form 5)

ER procedure. The students in Groups 1 and 2 participated in an extensive reading class for one academic year. There was no extensive reading during the first class, due to the orientation in April, nor in the two final classes of the spring and fall semesters, when the reading sections of the SLEP tests were administered. Thus, a total of 27 classes of ER and ER-and-Shadowing instruction were conducted over one academic year. Although in ER students choose the book on their own, the teacher selected some series of books for them, especially at the beginning of the course. Later, the teacher brought a variety of books for students, and they chose what they wanted to read in class. Students started reading very simple books which were much easier than their English level might suggest. The series they started reading, such as *Oxford Reading Tree* and *Longman Literacy Land*, were leveled for native English-speaking children (about K-3). The teacher added more book series each week, mostly leveled readers for children, such as *Step Into Reading*, *I Can Read Books*, *Time-to-Discover*, and *Rookie-Readers*. The level of the books increased gradually; therefore, the number and variety of books that the teacher brought into the classroom increased every week, as shown in Table 2. Before summer vacation, some students began to read thin, easy paperbacks with many pictures, such as the *Ricky Ricotta and the Mighty Robot* series as well as some graded readers. One student read one from the *Harry Potter* series, and another read *Frindle* by Andrew Clements. The teacher employed the three golden rules developed by Furukawa, Kawade and Sakai (2003) and Sakai and Kanda (2005).

The students were required to keep a record of their reading throughout the year. They had to write down the title of the book and a short comment in Japanese. Also, they reflected on their reading attitude or their feelings toward reading in a journal, and they wrote down titles of their favorite books. They were also required to keep a record of the number of books and words that they read in each class.

In addition to reading extensively in class, the students were strongly encouraged to read (and practice shadowing in Group 2) outside the classroom. Since the university has a well-equipped library, the teacher let the students select and read books from the library, in addition to drawing from the teacher's personal library. The teacher complimented the students who read books at the library. The teacher used stickers and *smileys* on students' journals to encourage them to read more and more.

Table 2. *Teaching plan for group 1 (ER)*

Week 1	Pretest (SLEP)
Week 2	Oxford Reading Tree (ORT) 0-1+
Week 3	ORT 2-3
Week 4	Step Into Reading (SIR) 1-2
Week 5	ORT 5-6, ICR 1-2, (e.g., *Frog and Toad*, *Mouse Tales*)
Week 6	Picture Books. (e.g., *Curious George*, *Mr. Putter and Tabby*)
Week 7	ORT 6-7, Longman Literacy Land (LLL), *Ricky Ricotta* series
Week 8	Ready-to-Read, *Nate the Great* series
Week 9	ORT 8-9, LLL 4-6, more "*Nate*" books
Week 10	Skyrider A, LLL 7-
Week 11	"*Nate*" books, Walker Stories
Week 12	Children's books in black and white. (e.g., *Rainbow Magic* series)
Week 13	Children's books in black and white. (e.g., *A to Z Mysteries* series, *Magic Tree House* series)
Week 14	Students chose the books they wanted to read.
Week 15	SLEP
Week 16-29	Students chose the books they wanted to read.
Week 30	SLEP

Shadowing procedure. At the beginning of the course, three materials, *Mouse Tales* and *Mouse Soup* by Arnold Lobel (2004) and *Nate the Great* by Marjorie Weinman Sharmat (2008), were selected. These were believed to be appropriate for the majority of the students. The teacher brought 30 portable CD players to the classroom, and after a demonstration by the teacher and practicing together as a class, each student practiced shadowing individually. In addition to reading the books described in Table 2, Group 2 (22 participants) practiced shadowing. About 20 to 30 minutes were used for shadowing, usually at the beginning of each class. In the fall semester, the teacher brought in various CDs accompanied by reading materials. Students practiced shadowing and read the books.

The best way to provide appropriate shadowing materials has not been established, and it is necessary to establish proven and precise procedures since in ER, there should be more

diverse materials geared towards students at a variety of levels in a wide range of genres. Also, the appropriate level of difficulty for shadowing has not yet been well examined. Shadowing is a way to help students recognize English prosody and follow the sound of English at a fast pace. It enhances and starts to define the sound loop and automaticity in the student's mind. When students are reading materials slowly, they can easily follow the story word by word, but shadowing is not just a reading comprehension activity. It might be crucial to use easy materials at the beginning in order to increase students' confidence, but they should move toward more difficult materials in order to appreciate the full benefits of this technique. Therefore, selecting and providing materials should be discussed between the teacher and students more often in shadowing class.

During shadowing practice, the teacher walked around and listened to the students as they practiced shadowing. Since the goal is to have students become familiar with shadowing and help them get into *the flow* of the activity, the teacher does not criticize students' performance at this stage; instead, the teacher only makes positive, motivating comments.

Results

The amount of reading students in Group 1 did during the entire course varied from 48,668 to 176,345 words, with an average of 73,646 words. Students in this group read between 72 and 203 books, with an average of 147 books. The number of words students in Group 2 read ranged from 25,636 to 562,394, with an average of 116,272; students in this group read between 82 to 310 books, with an average of 185 books. Figures 3 and 4 demonstrate that both the number of words and number of books read on average were greater in the ER-and-shadowing class than in the ER-only class.

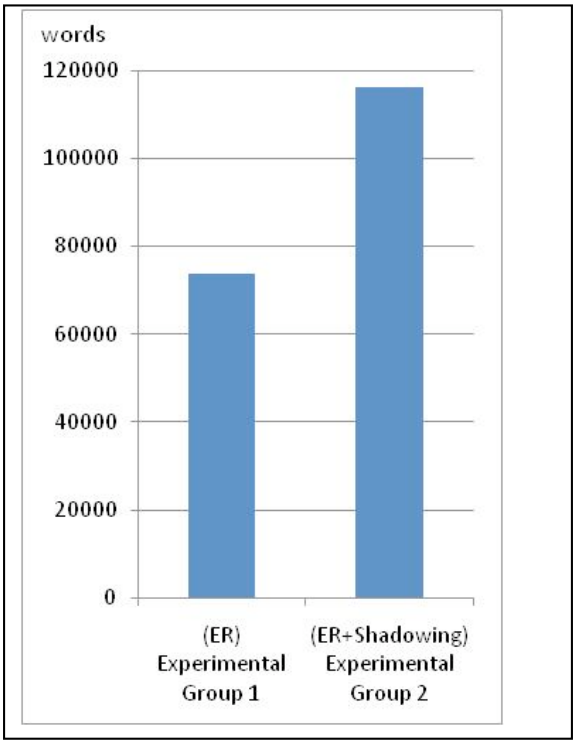

Figure 3. Number of words read

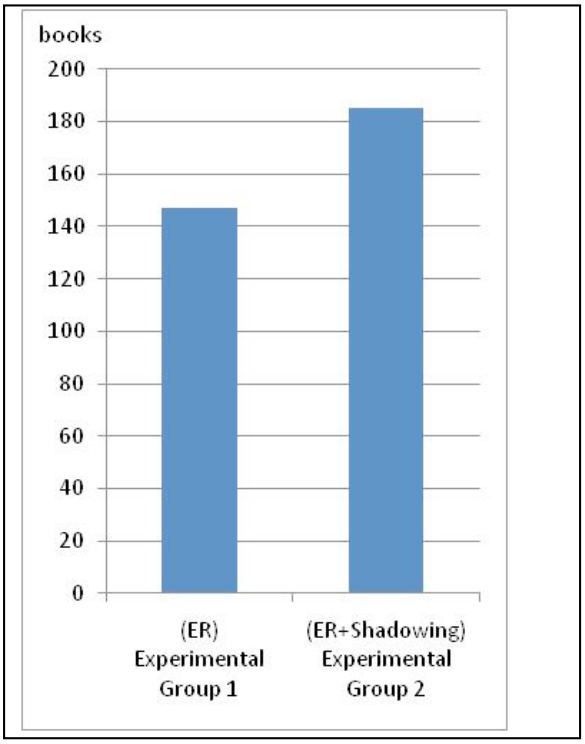

Figure 4. Number of books read

A one-way analysis of variance (ANOVA) was performed to evaluate the relationship among the four groups' scores on Test 1. Table 3 shows descriptive statistics of the four groups. In each case, the number of participants (*N*), mean (*M*) and standard deviation (*SD*) are given. The table also gives reliability estimates for the test. Cronbach's alpha (α) represents the percent of reliable or consistent variance in each group. For example, Cronbach's alpha suggests that Test 1 can be viewed as 72% reliable. The results of the one-way ANOVA are presented in Table 4. As the results indicate, there was no significant difference among the four groups, F (3, 85) = 1.53, p = .21, meaning that this study met the assumption that all groups were equal in terms of proficiency at the outset. Once this assumption was met, the posttests were administered after the instruction.

Table 3. *Descriptive statistics for pretest (α = .72)*

Group	*N*	*M*	*SD*
1 (ER)	20	29.30	8.27
2 (ER and Shadowing)	24	30.88	5.17
3 (Control 1)	21	32.76	6.25
4 (Control 2)	24	32.88	5.64
Total	89	31.51	6.41

Note. Full score is 71.

Table 4. *One-way ANOVA results for pretest*

Source	*SS*	*df*	*MS*	*F*	*p*
Between groups	184.99	3	61.66	1.53	0.21
Within groups	3425.26	85	4030.00		
Total	3610.25	88			

Note. $p > .05$

A two-way between-subjects analysis of variance was conducted to evaluate the effect of extensive reading and shadowing on reading comprehension tests. Descriptive statistics for all the tests and the results for two-way between-subjects ANOVA are displayed in Tables 5 and 6. Although the two control groups produced slightly better scores than the experimental groups, all four groups made steady progress. The test scores of the reading comprehension test (SLEP) were the dependent variables. The between-subjects factors were instruction methods with four levels (ER, ER and shadowing, control 1 and control 2). The test main effect and test x group interaction effect were assessed via the multivariate criterion of Wilks' lambda (Λ), which represents the ratio of error variance to total variance for each variate. The test main effect was significant, $\Lambda = .45$, F (2, 84) = 50.75, $p = .00$, eta-squared = .55; however, the test x group interaction effect was not significant, $\Lambda = .99$, F (6, 168) = .17, $p = .98$, eta-squared = .006, indicating that there is a significant difference among the three tests but no significant difference among the groups. The study set out to find the group difference but was unable to detect one.

Table 5. *Descriptive statistics for tests 1-3*

		Test 1		Test 2		Test 3	
Groups	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1 (ER)	20	29.30	8.27	31.60	6.48	36.65	7.39
2 (ER + Shadowing)	24	30.88	5.17	33.29	6.15	38.92	6.70
3 (Control 1)	21	32.76	6.25	34.95	7.68	40.52	8.95
4 (Control 2)	24	32.88	5.64	33.75	5.83	39.71	7.52
Total	89	31.51	6.41	33.43	6.53	39.00	7.64

Note. Full score is 71.

Table 6. *Results for two-way between-subjects ANOVA*

Source	*SS*	*df*	*MS*	*F*	*p*
Between subjects					
Group	453.22	3	151.07	1.56	0.21
Error	8221.31	85	96.72		
Within subjects					
Test	2676.26	2	1338.13	59.77	0.00*
Test x Group	21.75	6	3.63	0.16	0.99
Error	3805.74	170	22.39		
Total	15178.28	266			

Note. * $p < .05$

Post-hoc pair-wise comparisons were conducted to follow up on the significant effect of the

test. We controlled for family-wise error rate across these tests by making a Bonferroni adjustment. All three comparisons (Tests 1 and 2, Tests 1 and 3, and Tests 2 and 3) were statistically significant, meaning that all groups improved significantly in each test.

The plot shown in Figure 5 illustrates how the groups progressed among the three tests. The gains for each group (Test 3–Test 1) were as follows: ER = 7.35, ER & S = 8.04, C1 = 7.76 and C2 = 6.83. With all the groups showing gains on Tests 2 and 3, the ER and shadowing group demonstrated the most improvement.

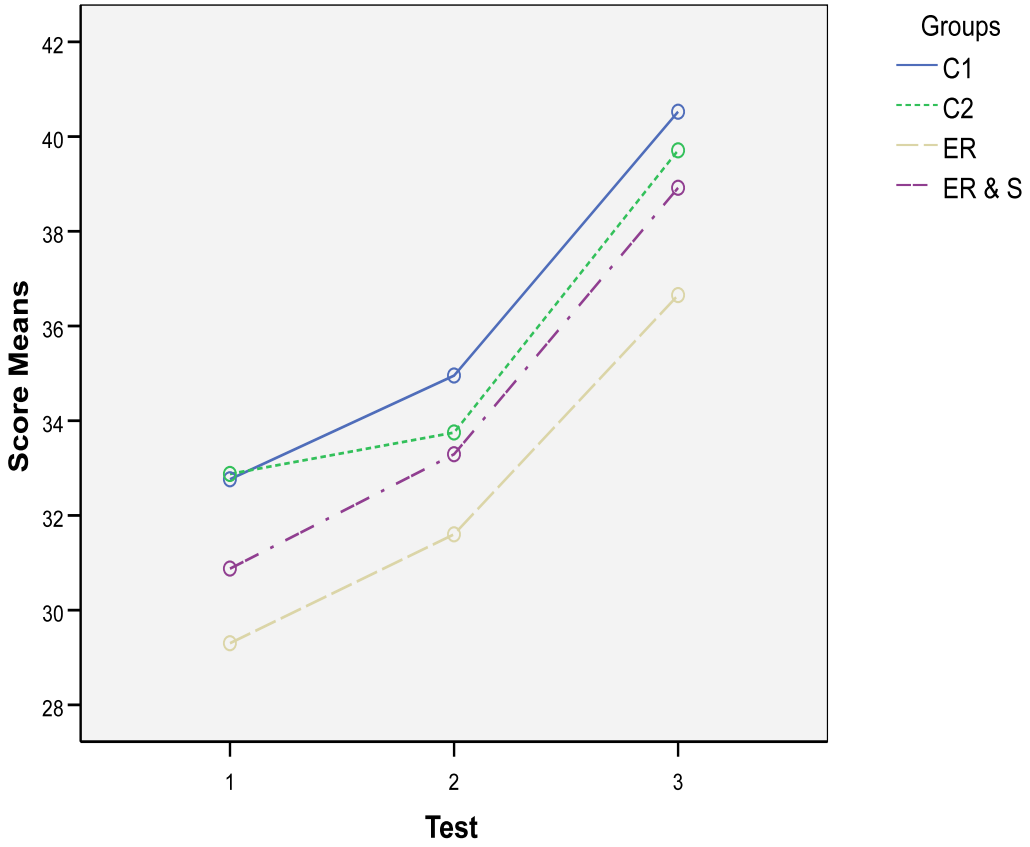

Figure 5. Plot of interaction of four groups' test scores.

Next, item facility (IF) analysis, which is "a statistic used to examine the percentage of students who correctly answer a given item" (Brown, 2005, p. 66), was performed to identify items flagged as misfits. IF analysis was conducted because the average scores of Tests 1 and 2 were less than 50% of the full score (35.5), meaning that the tests might not be of an appropriate level of difficulty for the participants. The cut-off IF percents for the items were set at the upper 90% and lower 20% levels. For instance, items with IF 98% or 15% were candidates for deletion.

Based on those IF indices, items 7, 15, 35, 37, 51 and 70 were deleted from both the pretest and posttest.

A two-way between-subjects analysis of variance was conducted with the remaining 65 items. The score main effect was significant, $\Lambda = .90$, F (1, 89) = 10.13, $p = .00$, eta-squared = .88; however, the score x group interaction effect was not significant, $\Lambda = 1.00$, F (3, 89) = .15, $p = .93$. Although the interaction effect was not significant, IF analysis examining eta-squared for SLEP with 71 items (.078) and 65 items (.88) functioned well on this test. One of the reasons for non-significance may be the small sample size, which may lead to small power (power = .08) in this study. Power analysis can be used for two reasons: planning and diagnosis. In this study, it was used for the latter. Murphey and Myors (2004) stated that it could be used "to determine whether a specific study has adequate power for specific purposes, or to identify the sort of effects that can be reliably detected in that study" (p. 17). According to Murphey and Myors (2004, p. 18), a power of .80 or above is usually judged to be adequate; therefore the required sample size to obtain a power of .80 was calculated using *G*Power 3* (Faul, Erdfelder, Lang, & Buchner, 2007). The input parameters were an effect size of .25, $p < .05$, 4 groups, 2 repetition, and correlation among rep measures of .50. The required sample size was 136, which is far more than the 89 participating in this study.

Discussion and Conclusion

The findings of this study can be summarized as follows. The first research question asked whether or not extensive reading is capable of improving students' reading comprehension as compared to control group. According to the posttest scores, extensive reading improved students' reading comprehension. Although a group difference could not be detected, posttest scores showed substantial improvement. Previous research (Krashen, 1982; Mason & Krashen, 1997; Robb & Susser, 1989; Yamashita, 2008) also supported the result.

The second research question asked whether or not shadowing could enhance the effects of extensive reading. When compared with the ER class, the ER-and-shadowing class showed more gains on posttest scores, indicating that shadowing seemed to enhance the effects of extensive reading. This was reflected through a comparison of posttest scores in which the gain of the ER-and-shadowing class was higher than that of the ER-only class. As shown in Table 5, however, the differences are less than one test score: ER & S = 2.41 vs. ER = 2.30 (Test 2 – Test 1) and ER & S = 8.04 vs. ER = 7.35 (Test 3 – Test 1). As indicated, the differences are only .11 and .69 respectively. Although the group difference could not be detected statistically in this study, further research on this perspective may produce more valuable results.

Since the results showed no difference between the four groups, it can also be concluded that ER, or ER plus shadowing, yield almost the same results as other conventional teaching methods (as done in the control groups). In other words, the ER program inside and outside of the classroom is at least as effective as conventional teaching, though we are currently unable to say that ER is superior or more effective than traditional teaching.

Six points in particular need to be addressed by future research. First, it appeared that the SLEP

was difficult for these participants. In fact, some participants had to be omitted from the study due to their inability to complete the test. An easier test, such as the A.C.E. (Assessment of Communicative English) developed by ELPA (Association for English Language Proficiency Assessment) or other tests, might be more accurate measures of the participants' proficiency.

Second, this study was only two semesters long. A longitudinal study may produce contrasting results. Fukada, Mishizawa, Nagaoka, and Yoshioka (2008) insisted that learners need to read more than 500,000 words in order to see the advantage of ER on the TOEIC, which would take more than three years on average.

Third, the sample size was too small to produce a power level of .80. In a repeated measures design, it is demanding and sometimes impossible to gather large numbers of participants; however, the power analysis indicated that the larger the sample, the better the results. Initially, 100 participants were present for Test 1, but as the semester proceeded, the number decreased for various reasons. Unfortunately, those who had taken two of the tests but could not be present at all three had to be eliminated.

Fourth, shadowing or the combination of ER-and-shadowing might motivate students more than ER-only instruction. Students in the ER-and-shadowing class read more, even though they had less time for reading in class due to the time they spent on shadowing. They utilized their time outside of class to read, and the effort was apparent to the teacher through her observations and during individual consultations in each class. It could be possible that doing shadowing in class motivated students to read more and pushed them to enter into the virtuous circle of the good reader, as Nuttall (2005) noted. Therefore, in the long run, the two types of instruction could produce contrasting results.

Fifth, according to the teacher's observations, in the ER classes, the students' attitudes toward English learning changed as students became more autonomous, though we were unable to compare this change with the control groups. The more students read, the more they tended to be conscious about their reading or reading attitude. Often, students wrote sentiments such as "the book is still beyond my reading ability, I'd like to try to read it later," "I feel confident reading, these days," "Now I know what kind of genre I like to read in English," or "Today, I happened to realize that I simply enjoyed reading and forgot that I was reading in ENGLISH, my least favorite subject" in their reading diaries. These comments suggest that students became more meta-cognitively conscious about their learning.

Finally, this study employed only a reading comprehension test, whereas shadowing actually required the students to read aloud; thus, shadowing could possibly enhance listening and speaking skills as well. The effects on listening and speaking skills need further research. As summarized in Ota (2007), the advantages of shadowing are as follows: First, students can familiarize themselves with the English phonological system due to extensive exposure to the language. Second, students will be able to develop speed by repeating sounds. And third, shadowing may help students concentrate on listening and help them feel a sense of achievement by being able to produce the original sounds.

Acknowledgments

We are grateful to two anonymous reviewers who provided very constructive comments and valuable feedback on the earlier versions of the paper. We would also like to thank Yo In'nami and Rie Koizumi for their insightful comments and advice. A paper based on this research was presented at the 6th Asia TEFL international conference, 2008, Bali, Indonesia.

Notes

1. Figures 1 and 2 adapted from Nuttall (2005). Reprinted with permission.
2. Takase did not specifically explain why raw scores were used, but we used raw scores in order to make it easier for readers to compare this study with hers.

References

Bamford, J., & Day, R. R. (2004). *Extensive reading activities for teaching language*. Cambridge: Cambridge University Press.

Brown, J. D. (2005). *Testing in language programs*. New York: McGraw-Hill.

Camiciottoli, B. C. (2001). Extensive reading in English: Habits and attitudes of a group of Italian university EFL students. *Journal of Research in Reading*, *24*(2), 135–153.

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

Déjean Le Féal, K. (1997). Simultaneous interpretation with "training wheels." *Meta*, *XLII*, *4*, 616–621.

Educational Testing Service. (2003). *Secondary level English proficiency test* [Technical supplement]. NJ: Author.

Elley, W. B., & Mangubhai, F. (1981). *The impact of a book flood in Fiji primary schools*. Wellington: New Zealand Council for Educational Research.

Faul, F., Erdfelder, E., Lang, A. -G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fukada, M., Nishizawa, H., Nagaoka, M., & Yoshioka, T. (2008). Toward the improvement of kosen students' English proficiency—Effects of SSS extensive reading. *The Council of College English Teachers Bulletin*, *27*, 1–8.

Furukawa, A., Kawade, M., & Sakai, K. (2003). *Ippai yomeba Shikkari minitsuku Kyokara Yomimasu Eigo 100mango* [The more you read, the better you become. Let's start reading one million English words today]. Tokyo. Nihon Jitsugyo Shuppansha.

Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh*. NJ: Pearson Prentice Hall.

Greenberg, D., Rodrigo, V., Berry, A., Brinck, T., & Joseph, H. (2006). Implementation of an extensive reading program with adult learners. *Adult Basic Education*, *16*(2), 81–97.

Hafiz, F. M., & Tudor, I. (1990). Graded readers as an input medium in L2 learning. *System*, *18*, 31–42.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study.

Canadian Modern Language Review, *61*(3), 355–382.
Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, *20*, 70–91.
Kadota, S. (2007). *Shadouingu to ondoku no kagaku* [Science of shadowning and reading aloud]. Tokyo: Cosmopier.
Kadota, S., & Tamai, K. (2004). *Ketteiban Eigo shadouingu* [Shadowing in English]. Tokyo: Cosmopier.
Krashen, S. D. (1982). *Principles and practice in second language acquisition*. New York: Prentice Hall.
Kweon, S. -O., & Kim, H. -R. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, *20*, 191–215.
Lai, F. -K. (1993). The effect of a summer reading course on reading and writing skills. *System*, *21*, 87–100.
Lobel, A. (2004). *Mouse tales audio collection* [Audiobook CD]. NY: Harper Children's Audio.
Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, *25*(1), 91–102.
Murphey, K. R., & Myors, B. (2004). *Statistical power analysis*. NJ: Lawrence Erlbaum Associates.
Muto, N. W. (2006). Why extensive reading? *Journal of School of Foreign Languages*, *Nagoya University of Foreign Studies*, *30*, 1–16.
Nakanishi, T. (2005). A book in question! Intertwining receptive and productive skills using graded readers. *The Language Teacher*, *29*(6), 35–36.
Nation, I. S. P. (1997). The language learning benefits of extensive reading. *Language Learning Online*, *21*(5).
Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford, UK: Macmillan Education.
Nye, P. W., & Fowler, C. A. (2003). Shadowing latency and imitation: The effect of familiarity with the phonetic patterning of English. *Journal of Phonetics*, *31*, 63–79.
Ota, E. (2007). The semantic processing encouragement in passage shadowing. *KATE Bulletin*, *21*, 25–36.
Robb, T. N., & Susser, B. (1989). Extensive reading vs. skills building in an EFL context. *Reading in a Foreign Language*, *5*, 239–251.
Sakai, K., & Kanda, M. (2005). *Koushitudeyomu eigo 100mango* [Reading one million words in the classroom]. Tokyo: Taishukan Shoten.
Takase, A. (2003). *The effects of extensive reading on the motivation of Japanese high school students*. Unpublished doctoral dissertation. Temple University, Tokyo.
Takase, A. (2004). Effects of eliminating some demotivating factors in reading English extensively. *JALT 2003 Conference Proceedings*. Tokyo: JALT
Takase, A. (2007a). Japanese high school students' motivation for extensive L2 reading. *Reading in a Foreign Language*, *19*, 1–18.
Takase, A. (2007b). Extensive reading in the Japanese high school setting. *The Language Teacher*, *31*(5), 7–10.
Ueda, A. (2005). Jirituteki gakushusha wo sodateru dokkai kurasu – Tadoku to purojekuto katudou wo chushintoshite [A language class employing ER and Projects–To have your students autonomous]. *Annual Report of Ibaraki University Education Center*, *9*, 41–50.
Waring, R., & Takahashi, S. (2000). *Guide to the 'why' and 'how' of using graded readers*. Tokyo: Oxford University Press, Japan.
Yamashita, J. (2004). Reading attitudes in L1 and L2, and their influence on L2 extensive reading.

Reading in a Foreign Language, *16*, 1–19.
Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, *36*(4), 661–672.

About the Authors

Takayuki Nakanishi teaches at Tokiwa University in Japan. His current research topics include investigations of Communicative Language Teaching, Willingness to Communicate, Language Testing and Extensive Reading. Email: tnakanis@hotmail.com

Atsuko Ueda received her BA in English and MA in international communication from Aoyama Gakuin University. After an eight year career as a materials and curriculum design for young learners of English, she started working as an English instructor in vocational schools in Tokyo. Her areas of interest include extensive listening and reading, teaching slow and young learners, learners' multiple intelligences, learning histories and motivation. She is an associate professor at Ibaraki University. Email: atzko@mx.ibaraki.ac.jp