The Effects of Reading-While-Listening and Listening-Before-Reading-While-Listening on Listening and Vocabulary

Boonyarit Tangkakarn
Language Institute, Thammasat University, Thailand, *ajboonyarit@yahoo.co.th*

Chanika Gampper
Language Institute, Thammasat University, Thailand, *chanika.g@litu.tu.ac.th*

This paper examined the comparative effects of Reading-While-Listening practice (RWL) and a proposed practice of Listening before Reading while Listening (LBRWL), both of which practices were inspired by the principles of extensive listening. Participants were 138 Thai Grade 8th students, from a school in the north of Bangkok, Thailand, divided into 3 groups, with 2 experimental groups (RWL and LBRWL groups) and a positive control group (high language proficiency group) as the top baseline. Two vocabulary tests and two listening tests (each of which had specific and general tests) were administered in three stages over 16 weeks (1 school term); before, after, and 10 weeks after the experiments. The results showed that participants, in both experimental groups, using 12 graded reader books with audio CDs, were found to gain in all four tests. The all delayed posttests indicated the language retentions. When compared among the three groups, the participants in LBRWL outperformed the RWL group. Additionally, there was no significant difference in performance between LBRWL and the positive control group, in each specific listening and vocabulary test. The questionnaire responses showed a positive attitude on both practices, but LBRWL showed a relatively stronger attitude.

Keywords: reading-while-listening, extensive listening, graded readers with audio, listening test, vocabulary test, listening practice

INTRODUCTION

Listening plays a great role in communication as Gilman and Moody (1984) reported in their research finding concerning how human spent time in one's life communicating, they found listening took up to 40-50%, while speaking, surprisingly, was only 25-30%, and reading around 11-16%, and writing about 9%. Listening must be of a great value, so much so that we spend a lot of time on it. The teaching of listening for L2 learners typically involved models and strategies like bottom-up, where teaching and learning of smaller to larger units, and top-down, where schema activation instruction was

employed (Flowerdew & Miller, 2005); listening strategies which included predicting/inferencing, deduction, etc. (Lynch, 2009); meta-cognitive strategy for listening (Wilson, 2008); intensive listening model when the teaching of pre-listening, while-listening, and post-listening was used (Richards & Burns, 2012). A majority of research on the teaching of listening models and strategies produced positive results, but still, most L2 learners (73%) found listening to be a particularly difficult skill to overcome (Chang, Wu, & Pang, 2013). One thing that was overlooked was the fact that learner needed to practice listening more. That is to say, listening is learned by simply listening and extensively.

Extensive listening (EL) can be any kind of listening that is pleasurable, extensive, self-selected by learners, comprehensible, and without exercises (Holden, 2005). Such principles of EL were adapted from the principles of Extensive Reading (ER). Despite being pleasurable and enjoyable listening, EL is quite taxing for L2 learners, as listening is hard because the nature of listening is filled with challenges, as it is real-time processing, fast speech rate, and possessing a variety of speech/accent/dialect (Renandya & Farrell, 2010). Therefore, the 'simultaneous listening and reading', or the Reading-While-Listening (RWL) was proposed to support listening practice (Stephens, 2010). Although the effects of RWL have already been known and quite well researched, the practice of RWL also met with criticisms. This study proposed a revised version of RWL by introducing 'Listening-Before-Reading-While Listening (LBRWL)' to improve language learners' listening ability.

RWL Practice

RWL has long been used in helping L1 learners in reading and in many cases for those with reading disability and the efficacy in helping the L1 learners in reading was confirmed by several studies (Bon et. al, 2001; Rasinski, 1990; Senechal & Cornell, 1993; Van der leij, 1981; Verlaan & Ortlieb, 2012; Winn et al., 2006).

In EFL and ESL reading, the efficacy of practicing of RWL was also confirmed (Antle, 2011, Billy, 2010; Blum et al., 1995; Elley & Mangubhai, 1981). Also, in EFL/ESL listening comprehension and vocabulary acquisition, RWL practice aided both, as supported by a number of research studies; for the effect on vocabulary acquisitions (Brown et al., 2008), and for that on listening comprehension (Chang, 2009; Webb & Chang, 2014).

However, Field's criticism (2008) towards RWL is worth taking closer attention, as he reasoned that RWL caused a 'divided attention' (p.56). He explained that between the sound and the text, L2 learners are more likely to pay more attention upon written text than upon spoken words. He further explained that although RWL helped to provide the word-sound relationship, the listening practice was lacked.

LBRWL Practice

The term and idea of LBRWL was actually stemmed from the statement made by Randall J. Lund (1991) who cautiously voiced the concern that by 'combining between listening and reading' (a practice of RWL) can be harmful to L2 learners, as "students will ignore listening and rely on reading" (p. 202). The author suggested the remedy that "Listening must be first because learners must learn to cope with the unknown" (p.202).

Similarly, Field (2008) supported the idea of 'first listening 'blind', then listening with the support of the text' (p. 56). The term 'blind' suggests the listening practice without support from visual text. By doing so, it prevents learners' pre-occupation with a text while trying to practice listening. Chang (2009), who researched extensively on the effect of RWL on listening comprehension, also suggested the idea of listening-only mode first "after a few attempts or failures, written scripts then provided to clarify their comprehension" (p. 662). Osada (2001) also agreed that providing written texts before listening to the text first was an inefficient way to practice listening. However, such suggestions by the afore-mentioned scholars above had not been proved empirically. Thus, this study aimed to compare the effects of two listening practices (namely, RWL and LBRWL) on listening comprehension and vocabulary knowledge, against a positive baseline control group to see how effective the two language interventions are.

METHOD

Participants

The participants were 138 eighth-grade students from 3 intact classes; selected purposively by the school administrator (the researcher had no control over the selection). The educational institution was a secondary school in the north of Bangkok, Thailand. The participants received formal EFL education since their primary school. These 3 classes were assigned into 2 experimental groups, with one group (50 participants) employed RWL practice, while the other group (50 participants) engaged in LBRWL practice, and 1 control group (38 participants). It must be noted that the control group is considered the highest proficiency group out of 3 groups. Thus, it is used as a positive baseline control group to which the experimental groups were compared. The head of English teachers initially informed that the language proficiencies of 3 groups were significantly different. This is a normal practice for typical Thai classrooms where students were placed in the room with the order from top academic class, to the lower academic class, right from the day students commenced their study. The three groups were typically taught with 4-skill language teaching as required by the language course book, mixed with grammar-translation approach. The researcher and two assistants performed the role as instructors as well as facilitators for the experimental classrooms. The participants, school head administrator and teachers gave consent for the study. A pilot study was conducted, using 30 volunteers who were excluded from the experiment, to learn the language intervention practice procedures, material use, listening device, test instruments, and survey questionnaire.

Materials and Listening Device

Twelve selected Dominoes series graded readers books from Oxford University Press (OUP), published from 2010 to 2013, were used in the study, with the rationales of (1) the suitability of the level for participants graded by word level from starter level at 250 word to level 2 at 700 word from wordlists ('headword' in OUP term) and by grammatical complexity, (2) illustrations which make story more interesting, (3) the end of chapter exercises to check progress of what are learned, (4) the suitability of speech narration rate ranging from 57 to 121 word per minute. Each graded reader book came with audio files narrating the stories. Twenty six sessions were used for the experiment. The audio mini mp3 player devices with micro-SD cards were used in the study.

Research Instruments and Data Collections

This study employed a quasi-experiment design. The instruments used to collect data were language tests and survey questionnaires. The language tests comprised of two listening tests and two vocabulary tests. Each of the listening and vocabulary tests has a specific and a general test to measure how much progress the participants achieved after interventions. The listening tests and vocabulary tests consisted of:

1) KET Listening test, a general listening test. This listening test was drawn from the past paper of authentic Key English Test (KET) for schools in book format, published by Cambridge University Press in 2014. KET is linked directly to the CEFR level of A2, which is suited for the students at Graded 8th, as stipulated by the Ministry of Education in Thailand (OBEC, 2015). The test consisted of 25 questions, separated into 5 parts, ranging from a short conversation to a long conversation, a long dialogue, and monologue. Test answering was in the format of multiple-choice, matching item, filling in the gaps, with answer keys.

2) GRs Listening test ('GRs' is short for 'Graded Readers'), considered a specific listening test was a 35 true/false questions. Participants answered questions while listening. The test was designed by the researcher specifically for this study, using two selected stories from Dominoes book series, titled 'The Real McCoy and Other Ghost Stories. The stories used were 'The Real McCoy' (pp. 1-5) and 'The Last Bus' (pp. 32-37), published by Oxford University Press in 2016. The test was piloted to ensure the reliability and both test reliabilities from Cronbach Alpha were 0.82, and 0.79, respectively.

3) GRs Vocabulary Test is the 80-questions translation test, a general vocabulary test. The vocabulary was selected from words listed in New General Service List (Browne, 2013). To receive a mark, the test takers were only required to translate a single meaning. No addition mark was given to additional meaning(s). The translation key was provided for 2 raters, and the inter-rater was also used to confirm the final score.

4) NVLT (New Vocabulary Levels Test) - a specific vocabulary test, consisting of 24 of 4-multiple choice questions. This is a vocabulary knowledge test, created by Stuart Mclean and Brandon Kramer (2015) to measure word knowledge from 1st 1000 word to 5th 1000 word. In this study, the 1st 1000 word level was used to match the vocabulary knowledge level of participants. The wordlists used were based on BNC (British Nation Corpus) and COCA (Corpus of Contemporary American English), developed and revised by Paul Nation (2012).

5) The 7-questions survey questionnaire was administered to all of the participants at the end of the immediate post-test stage to elicit their attitudes towards RWL and LBRWL practice directly related to listening comprehension and vocabulary knowledge.

Procedures

For RWL group, the practice steps are: (1) the participants listened to the narration while reading a text in the book to the story simultaneously; (2) The participants once again performed reading-while-listening. The rationale for the repetition was to control for the experiment. This made the time spent on the practices between groups equal, as

LBRWL group spent time listening first before reading-while-listening practice, while RWL group repeated the practice twice; (3) Participants answered questions only the first exercise (at the end of each chapter). The participants moved on to the next chapter and perform the same steps from step 1 to step 3, until completing the story, or when time was up (50 minutes per session).

For LBRWL group, the steps for LBRWL practice are; (1) the participants listened to the storytelling of a chapter without help from the text. (2) After finished listening to a chapter, the participants performed the reading-while-listening to the text of the same chapter again. This revealed the language they heard. Steps 3 for LBRWL is the same as steps 3 in RWL, until the story was finished, or when time was up (50 minutes per session).

The control group spent all 4 lessons per week learning English through traditional approach. All 3 groups received English language lessons 4 lessons a week for 16 weeks. Each lesson lasted for one hour. RWL and LBRWL groups took on average of 2 lessons a week and the remaining 2 lessons were taught using traditional approach.

A brief experiment procedure can be described as follows. After participants took four pretests, the experiment began, which lasted for one academic term (16 weeks, in 2016). That is, the students from both groups practiced RWL and LBRWL. After the experiments ended, the immediate posttest was administered. Then, the survey questionnaire was conducted. Ten weeks later, (at the beginning of the next academic term), the delayed posttests were administered, to measure language retention.

Statistical Procedure and Analysis

For data analysis, the descriptive statistics was first used to explore and describe the general performance of RWL, LBRWL and Control group through 36 tests. In answering the first and second research question, statistical comparisons of listening and vocabulary tests (KET Listening, GRs Listening test, GRs Vocabulary test, NVLT) were conducted by using one-way Repeated ANOVA statistics to compare 3 "within-groups" language growth before and after the interventions, and 10 weeks from post-intervention. Then, one-way analysis of variance (one-way ANOVA) statistics was used to confirm the head teacher's claim that the three groups English language proficiencies were different and this claim has been confirmed in that the pre-testing scores across-groups were statistically significant different and were regarded as covariates which would affect posttest scores. Thus, analysis of covariance (ANCOVA) statistics was performed to compare the score differences among 3 groups, by controlling for the pre-test score. In answering the third research question, a survey questionnaire was conducted and the responses were tallied and descriptively explained. With predetermined criteria, the interpretations for average responses for 7 questions were made, ranging from 'strongly disagree', 'disagree', 'agree', to 'strongly agree'. SPSS for Windows version 19.0 was used for statistical and data analysis. The statistical confident interval in this study was 95%, at p-value < 0.05.

Note: For ANCOVA statistics, there were a number of assumptions that needed to be strictly met, such as 'normality assumption' whereby dependent variables were assessed with Kolmogorov-Smirnov's Test of Normality, 'Homogeneity of the regression of

slopes assumption' where there was a test to assess that the slopes are homogeneous. All these assumptions have been tested and they were no violation of assumptions.

FINDINGS

Descriptive Statistics

The descriptive statistics in Table 1 illustrated the results of 36 sets of language test. The descriptive statistical report was arranged in the order of KET Listening test result, GRs Listening, NVLT, and finally, the GRs Vocabulary test. Within each type of test, the result of each group was reported by group, starting from RWL, to LBRWL, to control group, respectively. Then, the results were reported by stages of the test, from pre-test, to immediately posttest, and, finally, the delayed post-test.

Table 1
Descriptive Statistics of Score Results from All 36 Listening and Vocabulary Tests

Tests, stages, and groups	N	Min	Max	Total	Mean	Standard deviation
KET Listening Pre RWL	49	1	11	25	5.71	2.354
KET Listening Post RWL	49	3	14	25	8.12	2.505
KET Listening Delayed RWL	49	3	18	25	7.41	2.715
KET Listening Pre LBRWL	43	2	15	25	8.07	2.658
KET Listening Post LBRWL	43	3	22	25	10.72	4.382
KET Listening Delayed LBRWL	43	3	20	25	10.63	3.505
KET Listening Pre Control	38	4	21	25	13.79	4.70
KET Listening Post Control	38	5	23	25	14.08	4.756
KET Listening Delayed Control	38	6	22	25	14.29	4.86
GRs Listening Pre RWL	50	6	26	35	16.10	4.253
GRs Listening Post RWL	50	12	29	35	18.38	3.979
GRs Listening Delayed RWL	50	10	30	35	18.02	4.736
GRs Listening Pre LBRWL	50	12	31	35	22.10	4.987
GRs Listening Post LBRWL	50	10	34	35	24.52	5.048
GRs Listening Delayed LBRWL	50	11	34	35	23.38	5.558
GRs Listening Pre Control	31	18	31	35	25.65	3.536
GRs Listening Post Control	31	15	33	35	26.32	3.591
GRs Listening Delayed Control	31	14	33	35	26.52	4.781
GRs Vocab Pre RWL	47	1	32	80	14.30	6.788
GRs Vocab Post RWL	47	2	35	80	17.57	8.752
GRs Vocab Delayed RWL	47	7	49	80	22.09	10.579
GRs Vocab Pre LBRWL	47	7	40	80	21.64	7.654
GRs Vocab Post LBRWL	47	9	49	80	26.85	8.983
GRs Vocab Delayed LBRWL	47	8	53	80	30.91	10.157
GRs Vocab Pre Control	34	13	41	80	25.44	6.528
GRs Vocab Post Control	34	18	52	80	30.68	7.749
GRs Vocab Delayed Control	34	16	50	80	30.26	8.656
NVLT Pre RWL	49	6	21	24	12.39	3.724
NVLT Post RWL	49	8	19	24	13.94	3.092
NVLT Delayed RWL	49	4	24	24	13.45	4.154
NVLT Pre LBRWL	50	7	23	24	15.32	3.377
NVLT Post LBRWL	50	12	23	24	17.40	2.770
NVLT Delayed LBRWL	50	11	23	24	17.26	3.049
NVLT Pre Control	38	12	23	24	18.29	2.680
NVLT Post Control	38	17	23	24	19.87	1.647
NVLT Delayed Control	38	11	24	24	19.63	2.945

In general, the score results from each group grew over time, but each group was noticeably exhibited unequalled pre-intervention test results. For this, it is necessary to account for pre-testing scores statistically to confirm and compare the actual growths across-groups.

Research Questions 1: To what degree does engaging in LBRWL and RWL practices affect the participants' listening comprehension?

Using one-way Repeated ANOVA statistics to compare the results of "within-groups" general vocabulary tests, the RWL group elicited statistically significant changes in KET Listening score results over time, $F(2, 96) = 17.663$, $p < 0.0005$, partial $w^2 = 0.269$, with immediate posttest result increasing from that of pre-test, but not delayed posttest result. Secondly, the LBRWL group elicited statistically significant changes in KET Listening results over time, $F(1.43, 60.15) = 19.529$, $p < 0.0005$, partial $w^2 = 0.317$, with immediate posttest result increasing from that of pre-test, but not delayed posttest result. Meanwhile, the Control group did not elicit statistically significant changes in KET Listening results over time, $F(1.367, 49.204) = 13.425$, $p < 0.438$, partial $w^2 = 0.266$. The following Table 2 below showed the test result comparisons in three different testing stages of general listening test (KET Listening), along with statistical results.

Table 2
Post Hoc: "Within-Groups" Comparisons of KET Listening Mean Scores

KET Listening	Pretest	Immediate posttest	Mean difference	p-value
RWL group	5.71	8.12	*2.408	<0.0005
LBRWL group	8.07	10.72	*2.651	<0.0005
Control group	13.79	14.08	0.291	1.000
KET Listening (Continued)	Immediate Posttest	Delayed posttest	Mean difference	p-value
RWL group	8.12	7.41	0.714	0.243
LBRWL group	10.72	10.63	0.093	1.000
Control group	14.08	14.29	0.211	1.000

For specific listening test (GRs Listening test), it was found that the RWL group elicited statistically significant changes in test results over time, $F(2, 98) = 5.661$, $p = 0.006$, partial $w^2 = 0.104$, with immediate posttest result increasing from that of pre-test, but not delayed posttest result. Similarly, the LBRWL group elicited statistically significant changes in test results, $F(2, 98) = 5.649$, $p = 0.005$, partial $w^2 = 0.103$, with immediate posttest result increasing from that of pre-test, but not delayed posttest result. Control group did not yield a statistically significant different changes in test results, $F(1.678, 50.351) = 0.693$, $p = 0.481$, partial $w^2 = 0.023$, with no significant increase in both immediate posttest and delayed posttest results. Table 3 showed post-hoc analysis of specific listening test (GRs Listening tests).

Table 3
Post Hoc: "Within-Groups" Comparisons of GRs Listening Test Mean Scores

GRs Listening test	Pretest	Immediate posttest	Mean difference	*p*-value
RWL group	16.10	18.38	*2.28	0.007
LBRWL group	22.10	24.52	*2.42	0.005
Control group	25.65	26.32	0.677	1.000
GRs Listening test (Continued)	Immediate Posttest	Delayed posttest	Mean difference	*p*-value
RWL group	18.38	18.02	0.36	1.000
LBRWL group	24.52	23.38	1.14	0.253
Control group	26.32	26.52	0.194	1.000

In all, RWL and LBRWL groups exhibited growth in language test after intervention, but not the Control group. None of the groups showed significant language gains via delayed tests. To compare the "among-groups" listening comprehension tests, an ANCOVA statistics were carried, by controlling for pre-test scores.

Table 4
Adjusted and Unadjusted Means of Two Post-Intervention Listening Test Results with Pre-Intervention Test Results as Covariates

KET Listening					GRs Listening test				
	Unadjusted		Adjusted			Unadjusted		Adjusted	
	N	M	SD	M	SE	N	M	SD	M	SE
RWL	49	8.12	2.505	10.02	0.520	50	18.38	3.979	20.03	.682
LBRWL	43	10.72	4.382	10.95	0.497	50	24.52	5.048	23.99	.585
Control	38	14.08	4.756	11.53	0.625	31	26.32	3.591	24.51	.833

A one-way ANCOVA was run to determine the effects of language interventions on KET Listening scores. After controlling for KET Listening pre-test results (regarded as covariates) by adjustment of pre-intervention test results, using Bonferroni adjustment (see Table 4), Post-intervention KET Listening results among 3 different groups did not yield statistically significantly different results, meaning that all group made statistically the same gain, $F(2, 126) = 1.540$, $p = 0.218$, partial η2 = 0.024.

Whereas in the case of GRs Listening test, the two experimental groups and one control group yielded an interesting language gain. After controlling for covariates (pre-test scores) by adjusting for pre-intervention test results, using Bonferroni adjustment (see Table 4), post-intervention GRs Listening test results among 3 different group yielded statistically significantly different results, $F(2, 127) = 9.759$, $p < 0.0005$, partial $η^2 = 0.133$. To illustrate in more details, by comparing among-group GRs Listening posttest results, the LBRWL group result was better than that of RWL group ($M_{diff} = 3.958$, $p < 0.0005$) and the Control group outperformed RWL group ($M_{diff} = 4.471$, $p = 0.001$), but there was no statistically significant difference between LBRWL group and that of the Control group ($M_{diff} = 0.513$, $p = 1.000$). To sum up, post-intervention KET Listening results among 3 different groups did not yield any statistically significantly different, whereas post-intervention GRs Listening test results yielded statistically significant different. Both LBRWL and the Control group results were statistically significantly

better than that of RWL group, but the LBRWL group score result was not statistically significantly different from that of the Control group.

Research Question 2: To what degree does engaging in LBRWL and RWL practices affect the participants' vocabulary knowledge?

Using one-way Repeated ANOVA statistics to analyze by compare the results of "within-groups" general vocabulary tests, it was found that the RWL group elicited statistically significant changes in GRs Vocabulary score results over time, $F(1.562, 71.857) = 26.896$, $p < 0.0005$, partial $w^2 = 0.369$, with immediate posttest result increasing from that of pre-test, and also delayed posttest result was found to have increased from immediate posttest results.

Table 5
Post-Hoc: "Within-Groups" Comparison of GRs Vocabulary Test Mean Scores

GRs Vocabulary test	Pretest	Immediate Posttest	Mean difference	*p*-value
RWL group	14.30	17.57	*3.277	<0.0005
LBRWL group	21.64	26.85	*5.213	<0.0005
Control group	25.44	30.68	*5.235	<0.0005
(Continued)	Immediate Posttest	Delayed Posttest	Mean difference	*p*-value
RWL group	17.57	22.09	*4.511	<0.0005
LBRWL group	26.85	30.91	*4.064	<0.0005
Control group	30.68	30.26	0.412	1.000

Secondly, the LBRWL group elicited statistically significant changes in GRs Vocabulary score result over time, $F(1.631, 72.879) = 65.415$, $p < 0.0005$, partial $w^2 = 0.587$, with immediate posttest result increasing from that of pre-test, and also the delayed posttest result which was also found to have increased from immediate posttest results. Lastly, for the Control group elicited statistically significant changes in GRs Vocabulary score results over time, was statistically significantly different, $F(2, 66) = 19.772$, $p < 0.0005$, partial $w^2 = 0.375$, with immediate posttest result increasing from that of pre-test, but, however, not the delayed posttest result. From Table 5, one interesting note is that the mean differences for both RWL and LBRWL groups between immediate and delayed posttest of GRs Vocabulary test were found to be statistically significant. This means that participants performed better despite having no language intervention during the ten weeks period. The researcher later found that some participants studied the vocabulary list by taking the photo of the test, thinking that the experiment was over, which, in fact, it was not. They felt that many of the vocabulary listed in the test were worth studying for their benefit. This resulted in a better performance in the delayed test.

In comparing the effect of language intervention towards the results of specific vocabulary test, it was found that, firstly, the RWL group elicited statistically significant changes in NVLT score results over time, $F(2, 94) = 8.487$, $p < 0.0005$, partial $w^2 = 0.153$, with immediate posttest result increasing from that of pre-test, but, not the delayed posttest result. Secondly, the LBRWL practice group also performed statistically significantly different over time, $F(1.40, 55.158) = 24.574$, $p < 0.0005$,

partial $w^2 = 0.334$, with immediate posttest result increasing from that of pre-test, but, not the delayed posttest result.

Lastly, the Control group elicited statistically significant changes in NVLT score results over time, $F(1.668, 61.713) = 8.934$, $p = 0.001$, partial $w^2 = 0.194$, with immediate posttest result increasing from that of pre-test, but, not the delayed posttest result. Table 6 exhibited post-hoc analysis of specific vocabulary test (NVLT).

Table 6
Post Hoc Analysis: "Within-Groups" Comparison of NVLT Mean Scores

NVLT	Pretest	Immediate Posttest	Mean difference	*p*-value
RWL group	12.40	13.94	*1.542	<0.0005
LBRWL group	15.32	17.40	*2.08	<0.0005
Control group	18.29	19.87	*1.579	<0.0005
NVLT (Continued)	Immediate Posttest	Delayed Posttest	Mean difference	*p*-value
RWL group	13.94	13.45	0.521	0.596
LBRWL group	17.40	17.26	0.140	0.680
Control group	19.87	19.63	0.237	1.000

In all, all groups showed significant language growth after intervention. However, none of the immediate posttest and delayed posttest scores in both GRs Vocabulary and NVLT were significantly different, except for the GRs Vocabulary immediate and delayed posttest for both RWL and LBRWL groups. Next, to compare "among-groups" language test performance of two vocabulary tests, the ANCOVA statistic was used, by controlling for both vocabulary pre-test scores.

Table 7
Adjusted and Unadjusted Means of 2 Post-Intervention Vocabulary Tests with Pre-Intervention as Covariates

| | GRs Vocabulary Test | | | | | NVLT | | | | |
| | | Unadjusted | | Adjusted | | | Unadjusted | | Adjusted | |
	N	M	SD	M	SE	N	M	SD	M	SE
RWL	47	17.57	8.752	23.26	0.788	49	13.94	3.594	15.39	.310
LBRWL	47	26.85	8.983	25.15	0.715	50	17.40	2.770	17.28	.276
Control	34	30.68	7.749	25.15	0.897	38	19.87	1.647	18.15	.355

From Table 7 above, for GRs Vocabulary test results, after controlling for a covariate (pre-test scores) by adjusting for pre-intervention test scores using Bonferroni adjustment, post-intervention GRs Vocabulary test results yielded no statistically significant difference among groups after the language intervention, $F(2, 124) = 1.652$, $p = 0.196$, partial $\eta2 = 0.026$. Post hoc analysis showed that "among-groups" the GRs Vocabulary posttest results rendered no statistically difference between LBRWL and RWL groups ($M_{diff} = 1.889$, $p = 0.263$) and no statistically difference between LBRWL and Control groups ($M_{diff} = 0.002$, $p = 1.000$), and also, similarly, there was no statistically significant difference in GRs Vocabulary posttest results between RWL and the Control groups ($M_{diff} = 1.887$, $p = 0.437$).

Whereas in the case of NVLT post-test results, after controlling for the NVLT pre-test results (covariates) through Bonferroni adjustment as shown in Table 7, post-intervention NVLT test results yielded statistically significant difference "among-groups", after the 16-weeks long language intervention, $F(2, 133) = 15.937$, $p < 0.0005$, partial $\eta2 = 0.193$. Post hoc analysis showed that among group differences the NVLT posttest results rendered a statistically difference between LBRWL and RWL groups, where LBRWL group outperformed RWL group ($M_{diff} = 1.887$, $p < 0.0005$), and also there was a statistically significant difference in NVLT posttest results between RWL and the Control groups, where the Control group outperformed RWL group ($M_{diff} = 2.761$, $p < 0.0005$), but there was no statistically difference between LBRWL and Control groups ($M_{diff} = 0.874$, $p = 0.156$).

To sum up, post-intervention GRs vocabulary test results among three groups did not yield statistically significantly different, whereas post-intervention NVLT results among groups yielded statistically significant different post-test results. Both LBRWL and the Control groups scored statistically significantly higher than the RWL group, but LBRWL did not score statistically significantly different from the Control group.

Research Question 3: What are the attitudes of participants in each experimental group towards the practice they received (LBRWL and RWL)?

The survey questionnaire consisting of 7 questions was conducted, in which 47 participants from RWL and 50 participants from LBRWL groups answered the survey.

Q1: RWL/LBRWL practice helped to improve my listening ability.
Q2: RWL/LBRWL practice helped me to understand the gist when listening.
Q3: RWL/LBRWL practice helped to improve my English vocabulary knowledge.
Q4: RWL/LBRWL practice helped me to memorize vocabulary better.
Q5: RWL/LBRWL practice made me feel more comfortable in English listening.
Q6: RWL/LBRWL practice increased my confidence in English listening.
Q7: RWL/LBRWL practice helped me to understand how words are pronounced.

Table 8
Questionnaire Responses from Participants (n=47) in RWL Group

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	M	SD
1	0	5 (10.6%)	34 (72.3%)	8 (17.0%)	3.06	0.528
2	0	11 (23.9%)	31 (67.4%)	4 (8.7%)	2.85	0.556
3	2 (4.3%)	4 (8.5%)	32 (68.1%)	9 (19.1%)	3.02	0.675
4	1 (2.1%)	7 (14.9%)	26 (55.3%)	13 (27.7%)	3.09	0.717
5	1 (2.2%)	15 (32.6%)	22 (47.8%)	8 (17.4%)	2.80	0.749
6	1 (2.1%)	12 (25.5%)	27 (57.4%)	7 (14.9%)	2.85	0.691
7	0	7 (14.9%)	30 (63.8%)	10 (21.3%)	3.06	0.604

Table 9
Questionnaire Responses from Participants (n=50) in LBRWL Group

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	M	SD
1	0	2 (4.0%)	33 (66.0%)	15 (30.0%)	3.26	0.527
2	0	9 (18.0%)	39 (78.0%)	2 (4.0%)	2.86	0.452
3	0	3 (6.1%)	26 (53.1%)	20 (40.8%)	3.35	0.597
4	0	5 (10.0%)	29 (58.0%)	16 (32.0%)	3.22	0.616
5	0	12 (24%)	27 (54.0%)	11 (22.0%)	2.98	0.685
6	0	6 (12.0%)	33 (66.0%)	11 (22.0%)	3.1	0.580
7	0	3 (6.0%)	35 (70.0%)	12 (24.0%)	3.18	0.523

From Table 8 and 9, the 4-point Likert scale measured participants' attitudes, from 'strongly disagree' to 'strongly agree'. Noted that 3 participants from RWL group were absented during the survey questionnaire session and one respondent from RWL group missed answering question 2 and 5. The interpretation through numerical criteria for average response values was; 1 - 1.8, strongly disagree; 1.81 - 2.60, disagree, 2.61 - 3.2, agree; 3.21 - 4, strongly agree (Dornyei, 2010). On the basis of this criteria, for questions 1, 3, 4, LBRWL group responded, on average, with 'strongly agree', while RWL group with 'agree'. For other questions, both group responded with 'agree'. However, in general, the LBRWL group voiced positive opinion than the RWL group.

DISCUSSION

All participants from RWL and LBRWL gained in four different pre-and-post language tests (see Tables 2, 3, 5, and 6), with 2 tests measuring listening comprehension and other 2 tests on vocabulary knowledge. The Control group, however, did not make a statistically significant gain in both listening tests, but made statistically significant gains in both vocabulary tests. The no-gain in listening comprehension tests shows that the mean differences between tests are statistically insignificant, but the differences are still noticeable. Based on the comparative study of effect sizes of within-group pre-and-posttest results in both listening tests, the Control group did produce gains but are not large enough as compared to the other two groups. As this group possessed the highest level of English proficiency than the other two, they should, of course, make a gain after a term of study, whether a small or large gain. In addition, the listening tests may not suit the level of participants' language proficiency. A more advance test may be needed to gauge the difference. Traditional English language teaching approach, without doubt, should at least contribute to some degree of language growth.

Alternatively, the fact that the Control group was able to gain in vocabulary test suggested that both vocabulary tests could measure all group very well. In addition, the test results also show that the language education at school helped participants learn new words from the in-class language teaching. The four-skill student's course book and the grammar-translation method used in the teaching also emphasized vocabulary teaching to help students learning new vocabulary which helped the test gains. For the delayed post-test results, the statistical data confirmed that all three groups exhibited no gains or drops in performance after 10 weeks from the end of the experiments. Participants were able to maintain language ability in both listening comprehension and vocabulary knowledge.

For general language tests (KET Listening test, and GRs vocabulary test), when compared the language performance "across-groups", there were no significant mean differences. The reason was that the general test was far less sensitivity measurement to gauge the progress of listening comprehension. However, with specific language test (GRs Listening test, and NVLT), when compared the language performance "across-groups", LBRWL group outperformed RWL group in both tests, but LBRWL group performed no significant difference from the Control group (a positive baseline group). This suggests that LBRWL group, which was outperformed by the Control group in pre-tests, had improved to the level comparable with the Control group, after the experiment, while the RWL did not, even after pre-test scores were already controlled for using ANCOVA statistics. This is because the specific language test is the far more sensitive measurement to gauge the language progress than the general test.

In terms of participants' attitudes towards the practices, the survey results were overwhelmingly clear that majority of the participants seemed to have positive views towards the two listening practices, but LBRWL group showed more positively attitude than the RWL group.

The theoretical explanation of the effect of LBRWL practices comprises two scientific rationales, as to how the learners were linguistically supported. First, LBRWL practice seemed to benefit from Working Memory, which refers to a brain system that provides short-term memory storage and manipulation of the information necessary for complex cognitive tasks; such as language comprehension, learning, and reasoning (Baddeley, 1992). It was hypothesized that while RWL practice caused cognitive overload by visual and aural tasks, the LBRWL practice allowed a more balance for working memory to process and store without becoming too overloaded before the next round of aural and visual inputs were used. Second, the Noticing Hypothesis proposed by Schmidt (1990) that not all inputs are equal and only input which is noticed will be cognitively processed. Learners who performed listening-only first would be noticing some difficult sounds, which would later be revealed through texts.

Limitations and Suggestions

As random sample selection and assignment were not used, the study, therefore, could not claim for a generalization of results, and thus, it is considered a quasi-experimental study (Nunan & Bailey, 2009). Larger sample size in the future study would be ideal to facilitate generalizability of the study to larger population. The use of Meta-analysis by accumulating multiple small scale experiments is also possible to claim such generalizability of the experiment, and it is important that the experiment should include a control group, so that the comparative effect can be compared. To strengthen the result, future study should allow more study time of at least an academic year, not just one term (Nakanishi, 2015).

Teaching Implications

LBRWL can be applied in a variety of ways; such as, books, smart devices (phones and tablets), personal computers owned by most people of all ages. With the advent of digital technology by converging the Graded Reader e-books and digital audios, this has made the LBRWL practice even more practical. In addition, the practice of closed-

caption video viewing, which proved to be effective tool for language learning (Winke et. al., 2010), can also benefit the LBRWL practice by letting learners watch and listen without the English captions first and then re-watch the same section with captions.

There are suggested 9 guiding principles for implementation of LBRWL practice in the classroom, which were derived from this finding and also inspired by EL principles suggested by Holden (2005) and Renandya and Jacobs (2016). These guiding principles can be adjusted to suit the learning conditions. Such principles are as follows;

Principle 1, listeners should be allowed to choose books by themselves. Teacher should train and guide them at the initial stage. Principle 2, materials should be graded at the right level. Principle 3, LBRWL should be practiced first at the early stage of practice listening, until listeners are more proficient, before moving on to pure Extensive Listening when they become more advanced. Principle 4, LBRWL should be practiced in quantity - the more, the better. This is similar to the shared reading by Elley and Mangubhai (1983). Principle 5, from this study, it was found that by telling L2 learners the objective and the benefit of the practice, learners are motivated to practice in more enthusiastic way towards goal; such idea was also supported by Renandya and Jacob (2016). Principle 6, teachers should provide learners with continuous linguistic support and encouragements, instead of letting them struggle by themselves and feel demotivated. Principle 7, learners should be provided with post-practice activity, for language learning reinforcement, which also provides "linked skill" Nation and Newton (2009). Principle 8, implementing LBRWL in classroom does not need to be costly, as graded reader books are readily available with CDs. If there is no audio CD available, teachers can use read-aloud method to the students (Renandya & Farrell, 2010). Principle 9, teachers should aim at being able to lead L2 listeners from an in-class listening practice to the out-of-class practice.

CONCLUSION

This research finding confirms that both RWL and LBRWL practices are beneficial to the L2 learners in the area of both listening comprehension and vocabulary acquisition. However, it was found that the latter practice appears to be effective to the former. Therefore, this finding not only contributes to the existing knowledge in the field, but also advocates LBRWL practice, as suggested by Field, 2008, and Lund, 2001, to be used widely and regularly in language classroom.

REFERENCES

Antle, J. B. (2011). Extensive listening and how it affects reading speed. *The Language Teacher*, *35*(5), 3-5.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

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

Blum, I., Koskinen, P., Tennant, N., Parker, M., Straub, M., & Curry, C. (1995). Using audiotaped books to extend classroom literacy instruction into the homes of second-language learners. *Journal of Reading Behavior, 27*(4), 535-563.

Bon, W., Boksebel, L., Fontfreide, T., & Hurk, A. (1991). A Comparison of three methods of reading-while-listening. *Journal of Learning Disabilities*, *24*(8), 471-476.

Browne, C. (2013). New general service list (NGSL): Celebrating 60 years of Vocabulary Learning. *The Language Teacher, 7*(34)*, 13-16.*

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language, 20*(2), 136-163.

Chang, A. C.-S. (2009). Gains to L2 listeners from reading while listening versus listening only in comprehending short stories. *System*, *37*, 652-663.

Chang, A. C-S, Wu, B.W-P & Pang J. C-L. (2013). Second language listening difficulties perceived by low-level learners. *Percep. & Motor Skills, 116*(2), 415-434.

Dornyei, Z., (2010). *Questionnaires in second language research: construction, administration, and processing*. with Tatsuya Taguchi. New York: NY, Routledge.

Elley, W. B., & Mangubhai, F. (1981). *The impact of a book flood in Fiji primary schools.* Wellington/Suva: New Zealand Council for Educational Research/Institute of Education at the University of the South Pacific.

Elley, W. B., & Mangubhai. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, *19*(1), 53-67.

Field, J. (2008). *Listening in language classroom*. Cambridge: Cambridge University.

Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge: Cambridge University.

Gilman, R. A., & Moody, R. L. M. (1984). What practitioners say about listening: research implications for the classroom. *Foreign Language Annals, 17*(4), 331-334.

Holden, W. R. (2005). *Extensive listening: A new approach to an old problem.* Bulletin of Faculty of Humanities, Toyama University.

Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *The Modern Language Journal, 75*(2), 196-204.

Lynch, T. (2009). *Teaching second language listening*. Oxford: Oxford University.

McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, *19*(2), 1-11.

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, *49*(1), 6-37.

Nation, I. S. P. (2012). *The BNC/COCA word family lists*. Retrieved from www.victoria.ac.nz/lals/about/staff/paul-nation.

Nation, P. I. S., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.

Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, MA: Heinle, Cengage Learning.

OBEC. (2015). *Handbook of Thailand English education based on Common European framework of reference for languages (CEFR)*. Office of Basic Education Commission (OBEC), Ministry of Education (MOE), Thailand. Retrieved from http://english.obec.go.th/english/2013/index.php/en/2012-08-08-10-26-5/74-cefr.

Osada, N. (2001). What strategy do less proficient learners employ in listening comprehension? A reappraisal of bottom-up and top-down processing. *Journal of Pan-Pacific Association of Applied Linguistics*, *5*, 73-90.

Rasinski, T. (1990). The effects of repeated reading and repeated listening while reading. *The Journal of Educational Research, 83*(3), 147-151.

Renanadya, W. A., & Jacob, G. M. (2016). Extensive reading and listening in the L2 classroom. In W. A. Renandya, & P. Handoyo (Eds.), *English language teaching today*: *Linking Theory and Practice,* (pp. 97-110). New York, NY: Routledge.

Renandya, W. A., & Farrell, T. S. C. (2011). 'Teacher, the tape is too fast!' Extensive Listening in ELT. *ELT Journal, 65*(1), 52-59.

Richards, J. C., & Burns, A. (2012) *Tips for teaching listening: A practical approac*h. NY: Pearson Longman.

Schmidt, R.W. (1990). The role of consciousness in second language learning, *Applied Linguistics*, *11*(2), 129–158.

Senechal, M., & Cornell, E. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, *28*(4), 361-374.

Stephens, M. (2010). The primacy of extensive reading. *ELT Journal*, *65*(3), 311-313.

Van der Leij, A. (1981). Remediation of reading-disabled children by presenting text simultaneously to eye and ear. *Bulletin of the Orton Society, 31,* 229-243.

Verlaan, W., & Ortlieb, E. (2012). Reading-while-listening: Improving struggling adolescent readers' comprehension through the use of digital audio recordings. In J. Cassidy, S. Grote-Garcia, E. Martinez, & R. Garcia (Eds.), *What's hot in literacy-2012 yearbook*, (pp. 30-36). San Antonio, TX: The University of the Incarnate Word.

Webb, S., & Chang, A. S. (2014). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667-686.

Wilson, J. (2008). *How to teaching listening*. Harlow, Essex: Pearson Longman.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1) 65-86.

Winn, B., Skinner, C., Oliver, R., Hale, A., & Ziegler, M. (2006). The effects of listening while reading and repeated reading on the reading fluency of adult learners. *Journal of Adolescent & Adult Literacy*, *50*(3), 196-205.