The Effect of Reading While Listening on TOEFL Gains[†]

Peter GOBEL *

Kyoto Sangyo University, Faculty of Cultural Studies *

The present study investigated the effect of Reading While Listening (RWL) on overall English comprehension among a group of EFL students. It was hypothesized that the amount of RWL would be correlated with gains on a standardized test of English (TOEFL ITP). The current study was based on the results of previous studies with extensive reading (ER), the idea that L2 readers acquire and access lexical items in a similar fashion to L1 readers, and that development of such item-specific phonemic awareness would be significantly aided by an audio accompaniment to silent reading. A total of 162 ESL learners participated in the study. After two semesters of treatment, results demonstrated a statistically significant gain in pre-TOEFL ITP scores. A multiple regression analysis was then performed to determine the effect of quantity of RWL on TOEFL gain scores. Results indicated the amount of RWL was a significant predictor of TOEFL ITP gain scores. Subsequent theoretical and pedagogical implications are discussed.

Keywords: Extensive listening, TOEFL, Gain scores, Reading while listening, Amount of reading

1. Introduction

1.1. Extensive Reading as a way into Reading While Listening

In input-poor language learning environments, the teacher has few choices regarding increasing the amount of input for students. Given the limited amount of classroom exposure to the target language, alternative methods and media are necessary to ensure that students are getting a reasonable amount of input, and that they can access this input both in and outside of class. Since a teacher or tutor is not always available outside of class, activities that foster input with a focus on fluency (rather than accuracy) may be emphasized. One such form of input is extensive reading. Extensive reading (ER) has been called, 'the single most effective way to improve language proficiency' (Maley, 2005: 354), and proponents of ER suggest that it leads to considerable learning gains in the areas of reading, writing, vocabulary learning, and overall proficiency (see Day & Bamford, 1998 for an overview).

Extensive reading is a form of reading instruction where the students read in quantity, with the purpose of gaining a general understanding of what is read. The goal of extensive reading is to help develop good reading habits while building

up and supporting student's knowledge of vocabulary and structure (Richards & Schmidt, 2002). There are various ways to implement ER in a course or curriculum, but Day and Bamford (2002) have suggested a set of basic guidelines:

1. The reading material is easy.
2. A variety of reading material on a wide range of topics must be available.
3. Learners choose what they want to read.
4. Learners read as much as possible.
5. The purpose of reading is usually related to pleasure, information and general understanding.
6. Reading is its own reward.
7. Reading speed is usually faster rather than slower.
8. Reading is individual and silent.
9. Teachers orient and guide their students.
10. The teacher is a role model of a reader.

The primary difference between ER and traditional reading tasks and materials is that ER requires a large amount of reading compared with traditional materials, where intensive reading and decoding are the focus of the reading exercise. The theoretical foundation for ER in the L2 classroom setting

came initially from the input hypothesis, which argues for comprehensible input being a necessary and sufficient condition for L2 acquisition (Krashen, 1993). Not only the kind of input, but also the amount has been suggested as a key component in language acquisition, with Ellis (2005) stressing the importance of massive amounts of input for second language acquisition. Extensive reading using appropriate materials, such as graded readers, was suggested by Ellis as one of the easiest ways of to meet the demands of these two hypotheses. Research has reported positive effects of extensive reading on aspects of L2 ability such as reading speed (Bell, 2001), vocabulary (Horst, 2005), grammar (Yang, 2001), writing (Tsang, 1996), L2 linguistic ability (Yamashita, 2008), and general L2 proficiency (Mason & Krashen, 1997).

1.2. Reading While Listening

Aural input is essential to second language learning and acquisition, and research has shown that one main difference between successful and unsuccessful learners is the ability to use listening for language acquisition (Vandergrift, 1999). The amount and type of input that is required for a learner to acquire a second language (L2) differs with each learner, but it is widely accepted that L2 learners need a large amount of comprehensible aural input to help them acquire a language (Rost, 2006). As mentioned above, however, in an EFL setting, learners usually have limited opportunities for sufficient input because their native language dominates most of their in-class communication (e.g., classroom management, explanations, translation, etc.), the quality of the input differs from that of natural speech (teacher talk being the norm in most classroom settings), and most listening tasks are one-way in nature (as opposed to interactive). Consequently, when faced with an authentic listening task, L2 learners face a number of difficulties, e.g., fast speech, unfamiliar vocabulary, or inability to match the spoken form with the written form, all making comprehension of the aural input difficult. Therefore, some researchers suggest that L2 learners need external support while listening to aid comprehension, with these higher levels of comprehension motivating them to keep on listening (Goh, 1999; Chang & Read, 2006; Chang, 2008).

In the area of listening instruction, there are many studies that research ways to enhancing listening comprehension through the use of visual aids, mind maps, captions, and other

tools. These are used to build schema, aid in planning and preparation and thus, lessen the cognitive load on the listener. Most of these support tools have been found to aid listener comprehension, with an added side effect of having a positive motivating effect on the L2 learning. One common example of such support when using authentic materials would be captions, which provide L2 listeners with written support for negotiating the gap in knowledge and processing between the L1 and L2. Brought down to an even simpler level, tape scripts for textbook recordings serve the same purpose.

Both of the above examples are examples of students being asked to read while listening. Reading while listening is not a new technique, and has been used in L1 reading programs for decades, being especially effective for students with reading disabilities (Van Bon, Boksbeld, Friede, & Van den Hurk, 1991). Research undertaken to determine the benefits of reading while listening has largely been concerned with early readers in elementary school, in a native-speaker context. Reading stories to children is almost universally acknowledged as good pedagogy, and when it is done in an environment of shared reading or recreational reading, it also produces considerable gains in reading and listening skills (Senechal & Cornell, 1993).

Studies on the effectiveness of reading-while-listening for comprehension have claimed that since low-proficiency readers (both L2 and L1) tend to break sentences into small incoherent parts while they read, the sentence integrity is lost, and with it, meaning as well. With the teacher reading aloud, sentence and text integrity is preserved by pushing the readers to process larger semantic units, which in turn leads to better comprehension. As a result, when engaged in reading while listening, two things happen: learners may realize that a higher level of comprehension is possible; and learners reading process may change for the better (Amer, 1997; Dhaif, 1990).

As mentioned earlier, video captions are one of the most common forms of reading while listening. Many teachers would argue that captions are necessary for even higher-level students when watching video. Markham and McCarthy (2001) compared the impact of auditory input with L1 or L2 captions and found that listeners performed better with L1 captions, providing evidence for the argument that L2 listeners need a great deal of support when processing auditory input. However, one of the most common forms of

support in second language listening – reading while listening – has not received much attention. Perhaps this is because of its lack of authenticity. That is, in real life listening we seldom have a written form to refer to. Another reason may be the difficulty of providing listeners with aural scripts (although it is becoming increasingly easy to obtain scripts in this digital age). There have been a few recent studies that have pointed out the benefits of reading while listening, assuming it is helpful for L2 listening comprehension and development (Bell, 1998; Vandergrift, 2007; Chang, 2008), and vocabulary acquisition (Brown, Waring, & Donkaewbua, 2008). Vandergrift (2007) in particular, found that reading while listening created the aural–written verification stage which was particularly valuable to low-proficiency groups for developing auditory discrimination skills, and to high-proficiency groups for refined word recognition. Similarly, Osada (2001) suggested that matching aural text with a transcription of the text can help listeners develop an awareness of form–meaning relationships and word recognition skills. Finally, Chang (2008) found that reading while listening was preferred over listening only by a majority of student in his study, leading him to suggest that reading while listening has a positive effect on task completion and performance.

Despite support for aural–written verification during L2 learning, it is unclear whether reading while listening enhances listening skill as such. One clear difference between reading and listening is the presence or absence of a printed text. Spoken language is made up of many different sounds that are, by their very nature, fleeting and impermanent. In other words, the listener hears the sound, and the sound is gone, with no text to refer back to. In addition, in real-world listening, listeners have little or no control over input. With reading, on the other hand, readers can control the speed of information processing, and can refer back to the text whenever they feel the need.

The present study is concerned with the practice of reading while listening (RWL) and its effect on English language proficiency. Since it is difficult to determine exactly what a student is doing or focusing on when s/he is practicing RWL, it was decided that more general measures of proficiency, rather than specific measures (e.g., reading speed, vocabulary acquisition, phonemic encoding) should be the focus of the study. The study was set up to meet the twin goals of providing comprehensible input and amounts of input larger than students would normally be exposed to in a regular EFL course curriculum. As such, graded readers, with accompanying CDs, were chosen as the medium, and a self-access approach to RWL was chosen as the method.

The research questions are as follows:
1. What effect, if any, does RWL have on TOEFL ITP gain scores?
2. Is the amount of RWL correlated with TOEFL gains?
3. Does vocabulary level of the material have any correlation with TOEFL gains?

2. Materials and Methods

2.1. Participants

162 first-year university students in the Faculty of Cultural Studies took part in the RWL program designed to boost their reading and listening input. The participants were streamed into 8 classes based on a placement test given in April, and were enrolled in a mandatory year-long English listening course (ELT IV) as well as three other required English courses. TOEFL scores ranged from 310 to 500. See Table 1 for a description of the participants.

2.2. Reading while Listening materials

The materials for the project were chosen from the extensive reading materials available at the university library and in the university language laboratory. These graded reading materials have been divided into reading levels based on the number of head words used in each book. Headwords are words within a level-appropriate list available to authors or adapters to use freely. These head words are chosen based on their frequency in written English, and students at the appropriate level can be expected to be familiar with these words. In the case of this study, Level 1 books containing 200 head words, Level 2 books containing 300 head words,

Table 1. Description of Participants

Group	Pre-TOEFL Range	N
1	350-500	36
2	310-440	85
3	287-410	41

Table 2. Average words per book and number of headwords

Level	Head Word	Word
	Avg	*Avg*/ Book
Level 1	200-300	913
Level 2	300-400	1885
Level 3	600-700	3005
Level 4	1000-1200	5133

Level 3 books containing 600 head words, and Level 4 books containing 1,200 head words were chosen. Please note that as the level of the material goes up, so does the average number of words per book. For this reason, the number of words, rather than the number of books, was used when calculating amount of RWL in this study.

A large percent of the extensive reading materials come with accompanying CDs, in which the text is read aloud. The recordings are clear, but since these are readings of texts aimed at EFL learners, the reading speed, and certain aspects of pronunciation have been altered. Thus, the recordings cannot be considered to be 'native speech', but are adequate for the purposes of RWL.

In addition to the graded readers available in the library, online quizzes available on the MoodleReader, were also a component of the program. These online quizzes were created for the purpose of testing students' comprehension of the graded readers they had supposedly read. Students were supposed to complete a graded reader, log on to the MoodleReader, and take the quiz in a ten-minute time span. The quiz is then graded and the results saved on the MoodleReader. In addition to quiz scores, other information, such as number of books read, number of successful quizzes taken, and total number of word read are recorded on the student log.

2.3. Procedure

Students were assigned to graded reader levels based on their initial placement test scores, with the top 36 students starting at Level 3, the mid 85 students starting at Level 2, and the lower 41 students starting at Level 1 (Table 1). Students were told that they had to listen to a minimum of five graded readers over each semester (10 in total), and pass the on-line quiz for each of these books. This project would

account for 20% of each semester grade. If students read more, they would receive additional points, thus raising their semester grade. In addition, after successfully passing three quizzes, if students wished to practice RWL at a higher level, they were allowed to do so. Following an orientation on the MoodleReader program and the RWL procedure, students were left on their own to check out books, read while listening, and take the quizzes. Some students in each level also received in-class practice with RWL, but this was not controlled across classes, and therefore was not included as a variable in the study.

Every month or so the teachers in each class would check the student logs and encourage the students to complete their assignments. At the end of the year the researcher noted the total number of words each student had completed (based on the data in the MoodleReader files). It was felt that the number of words was a more accurate record of time spent on task than either number of successfully completed quizzes, or number of books (since book length varied greatly depending on the level). Finally, the students took a TOEFL ITP in December, and these scores were then compared with the April scores to calculate gain scores for all sections of the test and the overall score.

A MANCOVA using the preliminary TOEFL ITP scores as the covariate was performed in order to see if gain scores were significant. To see whether amount of RWL and difficulty of the material (level) were predictors of TOEFL gains, a multiple regression analysis was performed, using gain scores on the total as the dependent variable.

3. Results

3.1. MANCOVA

A one-way multiple analysis of covariance (MANCOVA) was conducted. The independent variable, amount of RWL included three levels: 12,000 to 9,999 words, 10,000 to 18,999 words, and 19,000 to 100,000 words. The other independent variable, material level included three levels as well: Level 1 (200 head words), Level 2 (300 head words), and Level 3 (600 head words). The dependent variable was the overall gain on the TOEFL ITP test from April to December administration, and the covariate was the initial total TOEFL score on the April administration. A preliminary analysis evaluating the homogeneity of slopes assumption indicated that the relationship between the covariate that

Table 3. Average TOEFL Gain by Group

Group	Average Gain	Std. Error	Lower Bound	Upper Bound
1	36.528	4.548	27.558	45.499
2	18.622	2.128	14.424	22.820
3	10.270	4.357	1.675	18.865

independent variables did not differ significantly as a function of the independent variables $F_{(2, 158)} = 3.658$, MSE = 465.745, p = .28 The MANCOVA was significant $F_{(1, 158)} = 56.49$, MSE = 478.872, p < .00. The strength of relationship between the amount of RWL, reading level, and the dependent variable was reasonably strong, as assessed by a partial eta2, with the RWL factor accounting for 30% of the variance of the dependent variable, and the reading level accounting for 18% of the variance.

The means of the TOEFL gains adjusted for initial differences were ordered as expected across the three RWL groups and the reading level groups (see Table 3).

3.2. Multiple Regression

A multiple regression analysis was conducted to evaluate how well the amount of RWL and material level predicted TOEFL gain scores. The predictors were amount of RWL and level of the material. The linear combination of the two predictors was significantly related to the TOEFL gain $F_{(2,160)} = 6.100$, p = .003 . The sample multiple correlation coefficient was .26, indicating that approximately 7% of the variance in TOEFL gains in the sample can be accounted for by the linear combination of the predictors.

4. Discussion

In the section below, the three research questions will be answered together with extended discussion of the results.

RQ1. What effect, if any, does RWL have on TOEFL ITP gain scores?

Based on the results of the study, the answer to the first question is that RWL has a positive effect on TOEFL scores. Using the April test scores as a covariate allowed the researcher to look at individual gains as compared with the amount of RWL, regardless of the initial TOEFL score of the student. The April covariate also eliminated the effect of preliminary test score as a confounding variable. With these problems out of the way, individual gains based were

compared with the amount of RWL and the comparison was found to be statistically significant. In what way the comparison was significant was the purpose of the second research question.

RQ2. Is the amount of RWL correlated with TOEFL gains?

The results of the multiple regression indicate that the amount of RWL is a solid predictor of TOEFL gains. In other words, with this sample, the more the students practiced RWL, the greater their gains. It must be noted, however, that the amount of RWL was quite small in comparison to the total number hours of instruction. This may explain why the RWL factor played such a small role in predicting TOEFL gains.

RQ3. Does vocabulary level of the material have any correlation with TOEFL gains?

In the case of the third research question, it seems that level of the material, which was directly connected to placement test scores had a minor effect on TOEFL gains. This can be interpreted in two ways: that the level of the vocabulary was directly related to the development of processing, or that the higher level texts allowed the students to incidentally learn vocabulary which helped them on the TOEFL. The former interpretation makes sense and is supported by a number of studies on vocabulary recognition and reading speed following treatment by RWL (see Askildson, 2008 for an overview).

As for the latter interpretation, while this is possible, the vocabulary level that the students are using for RWL is much lower than their theoretical amount of passive vocabulary. It would be hard to imagine that the students were actually learning a significant amount of new vocabulary through this practice. Rather, is suggested that the higher levels pushed students to process more, and faster, than the lower level texts. In addition, it is probable that the higher level texts also had a higher word per minute ratio than the lower level texts. This was not controlled in this study, but would have a great effect

on processing. This is not to suggest that higher level texts are better for students. Please bear in mind that all of the texts were at quite a low level (when considering head words) and thus, far removed from authentic texts and listening.

It must be noted that, unlike previous studies that looked at specific gains in reading speed, vocabulary acquisition, and other limited factors, this study looked only at TOEFL gains as a representative of overall comprehension improvement. The TOEFL was the only reliable test available to the researcher at the time of the study. The TOEFL is a test of overall comprehension ability, and as such the author did not feel it was sensitive enough to test separate constructs of proficiency such as vocabulary, reading speed, or listening comprehension.

5. Conclusion

This was a study that aimed to investigate the effect of RWL on gains in comprehension. It should be pointed out that the results of the study should be considered from a general listening instruction perspective. In other words, the purpose of RWL should not be to raise test scores, but to improve comprehension and explore new ways to offer aural input in an EFL setting. This study supports the idea that RWL is an effective tool in offering input to students in a reasonably entertaining fashion, giving them choices regarding what they listen to, as well as when and where.

As this is a pilot study, there are a number of limitations inherent in the design. First and foremost, although students were trained in RWL, since this was an out-of-class activity, there was no way of controlling whether the students performed the activity correctly. Quiz results were used to ensure that the students finished the material (although cheating on quizzes had been reported) but exactly how the students approached the material is unclear. In an attempt to clarify this, two classes were quizzed as to how they worked with the material, and both classes reported that they felt they had completed the RWL assignments properly. However, activities performed outside of the classroom are very difficult to monitor and control, and this limitation is inherent in all such studies, including those done on extensive reading. Another limitation was the use of the TOEFL gain scores as a measure of improvement. As mentioned earlier, the TOEFL is a general measure of English comprehension, and a crude one at best. There have been numerous studies criticizing the TOEFL test and its inability to measure the constructs it purports to measure (Freedle & Kostin, 1999). Future research might want to consider more sensitive instruments, and instruments that measure more select constructs such as processing speed, reading speed, etc.

Still another limitation was that this study used non-communicative speech as the treatment method, and student gains were assessed through the TOEFL, a non-productive measure of language. Future research could investigate more authentic aural input such as two-way listening tasks, to better understand what other language elements of listening can be acquired.

Finally, the amount of RWL in this study was admittedly small. A rough calculation of the number of hours of RWL per student worked out to an average of four hours each semester (17,400 words). This can hardly be called 'extensive listening', since there was little more than an average of sixteen extra minutes of input a week (assuming the students practiced RWL steadily, every week). On the other hand, the fact that this small amount of listening seemed to have a positive effect is encouraging, and future studies should investigate larger amounts of RWL with more diverse populations.

One other point of consideration was that of motivation to practice RWL. With Japanese students, getting them to complete out of class tasks is always a challenge (one possible explanation for the low number of hours RWL). With the greatest amount of words accessed in RWL being 99,000 we can see a clear distinction between those who practiced RWL a lot, and the average. This suggests the possible impact of motivation as an intervening variable, and this should be dealt with in future studies.

REFERENCES

Askildson, L. R. (2008) *"The effects of reading while listening: Developing L2 reading via phonological recoding"*, Unpublished manuscript.

Amer, A. (1997) *"The effect of the teacher's reading aloud on the reading comprehension of EFL students"*, ELT Journal, 51, pp. 43–pp. 47.

Bell, T. (1998) *"Extensive reading: why and how?"*, The Internet TESL Journal, 6.

Bell, T. (2001) *"Extensive reading: speed and comprehension"*, The Reading Matrix, 1(1).

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, pp. 136–pp. 163.

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

Chang, C.-S., Read, J. (2006) *"The effects of listening support on the listening performance of EFL learners"*, TESOL Quarterly, 40, pp. 375–pp. 397.

Day, R., J. Bamford (1998) Extensive Reading in the Second Language Classroom, Cambridge: Cambridge University Press.

Day, R., Bamford, J. (2002) *"Top ten principles for teaching extensive reading"*, Reading in a Foreign Language 14, pp. 136–pp. 141.

Dhaif, H. (1990) *"Reading aloud for comprehension: A neglected teaching aid"*, Reading in a Foreign Language, 7, pp. 457–pp. 464.

Ellis, R. (2005) *"Principles of instructed language learning"*, System, 33, pp. 209–pp. 224.

Freedle, R., Kostin, I. (1999) *"Does the text matter in a multiple-choice test of Comprehension? The case for the construct validity of TOEFL's minitalks"*, Language Testing, 16(1), pp. 2-pp. 35.

Goh, C. (1999) *"How much do learners know about the factors that influence their listening comprehension"*, Hong Kong Journal of Applied Linguistics, 4, pp. 17–pp. 40.

Krashen, S. (1993) The Power of Reading: Insights from the Research, Libraries Unlimited, Englewood, CO.

Maley, A. (2005) *"Review of 'Extensive Reading Activities for the Second Language Classroom'"*, E LT Journal, 59, pp. 354–pp. 355.

Markham, P., Peter, L., McCarthy, T. (2001) *"The effects of native language vs. target language captions on foreign language students' DVD video comprehension"*, Foreign Language Annals, 34, pp. 439– pp. 445.

Mason, B., Krashen, S. (1997) *"Extensive reading in English as a foreign language"*, System, 25, pp. 91–pp. 102.

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

Richards, J.C., Schmidt, R. (2002) Longman Dictionary of Language Teaching and Applied Linguistics, third ed. Pearson Education, London.

Rost, M. (2006) *"Areas of research that influence L2 listening instruction"*, In: Uso´-Juan, Mart´ınez.-Flor, A. (Eds.), Current Trend in the Development and Teaching of the Four Language Skills, Mounton de Gruyter, Berlin, pp. 47–pp. 74.

Senechal, M., Cornell, E. (1993) *"Vocabulary acquisition through shared reading experiences"*, Reading Research Quarterly, 28, pp. 361–pp. 374.

Tsang, W. (1996) *"Comparing the effects of reading and writing on writing performance"*, Applied Linguistics, 17, pp. 210–pp. 233.

Van Bon, W. H. J., Boksebeld, L. M., Freide, T. A. M., Van den Hurk, A. J. M. (2001) *"A comparison of three methods of reading while listening"*, Journal of Learning Disabilities, 24, pp. 471- pp. 476.

Vandergrift, L. (1999) *"Facilitating second language listening comprehension: acquiring successful strategies"*, ELT Journal, 53, pp. 168–pp. 176.

Vandergrift, L. (2007) *"Recent developments in second and foreign language listening comprehension research"*, Language Teaching, 40, pp. 191–pp. 210.

Yamashita, J. (2008) *"Extensive reading and development of different aspects of L2 proficiency"*, System, 36, pp. 661-pp. 672.

Yang, A. (2001) *"Reading and the non-academic learner: A mystery solved"*, System, 29, pp. 451–pp. 466.

和文抄録

　本研究では、朗読音声を聴きながら英文を読む練習（Reading While Listening）が総合的な英語力にどのような影響を与えるかのかについて調査した。本研究は、多読に関する先行研究、および外国語学習者も母語を学んだ時と似た方法で語彙を習得するため、音声を聴きながら黙読することによって音素を識別する能力が伸びるという仮説に基づくものである。162人の学生が調査に参加し、1年間にわたって朗読音声を聴きながら英語を読む練習をした結果、統計的に有意なTOEFLスコアの伸びが見られた。また、重回帰分析の結果によると、音声を聴きながら読んだ英文の量とTOEFLスコアの伸びの間に相関関係があることもわかった。

キーワード： 多聴, TOEFL, スコアの伸び, 音声を聴きながら英文を読む練習, 多読量