Gains to L2 Learners from Extensive Listening: 
Listening Development, Vocabulary Acquisition and 
Perceptions of the Intervention

Ching-Shyang Anna Chang 
Department of Applied English  
Hsing-Wu University

Abstract

Building on evidence of the success of extensive reading in L2 learners’ 
language development, this study explores the effects of two listening types,  
extensive listening (EL, n = 31) versus intensive listening (IL, n = 24), on 55  
Chinese university students of English over a 26-week period using audio 
recordings of graded readers to explore students’ listening development and 
vocabulary acquisition. The EL and IL groups studied a total of 15 and three 
audio readers respectively. The test measurements used to gauge the two 
groups’ language gains involved a 100-item listening test and a vocabulary 
levels test. The test instruments were administered to the students before the 
treatment and were repeated when the treatment was finished. The results 
showed that in the post-intervention test the EL group significantly 
outperformed the IL group, while in the vocabulary levels tests, the IL group 
made a significant gain and the EL group did not. Two major factors are seen 
as contributing to listening improvement: students being able to control their 
own learning (monitoring their comprehension and seeking their own 
solutions), and interesting scenarios being used in the study material. 
Limitations and scope for future research are discussed.

Keywords: extensive listening, extensive reading, listening development, L2 
vocabulary development

Introduction

In recent years in second language (L2)1 teaching programs, the status of 
listening has undergone a substantial change. Listening is no longer considered just a 
skill to comprehend a target spoken language but an avenue to acquire linguistic 
knowledge (Richards, 2005; Rost, 2002). This change implies that developing good 
listening skills can facilitate L2 development. However, in a foreign language 
environment, comprehending the target spoken language without difficulty is by no 
means easy, and factors affecting listening difficulties have been widely discussed 
both in first language (L1) (Samuels, 1984; Watson & Smeltzer, 1984; Wolvin & 
Coakely, 1996) and second language listening (Anderson & Lynch, 1988; Boyle, 1984; 
takes place when listeners can infer what is said based on their background, and 
linguistic and contextual knowledge (Buck, 1995, 2001; Rost, 1990; Widdowson, 
1983). However, unlike written language, which tends to be more stable, variations in 
spoken linguistic features may occur from person to person or region to region (Biber, 
1988; Chafe, 1985). Accordingly, listeners face a number of challenges, such as
connected speech, fast speech rates, accents, transient information, or colloquial usages and slang, which seldom appear in formal L2 textbooks.

In a comparison of L1 and L2 listening in a learning situation, Rost (1994; 2006) found the opportunities for input to be one of the major factors making L2 listening more difficult than L1 listening. Thus, providing learners with helpful interventions that will encourage them to listen more closely and with heightened curiosity seems to be one of the major functions of language teachers (Rost, 2007). How, then, can language teachers enrich their students’ aural input? Building on the effectiveness of extensively reading graded books in improving L2 learners’ reading skills (see Day & Bamford, 1998; Krashen, 2004 for comprehensive reviews), a few scholars, such as Brown (2007), Chang (2009, 2011), Renandya (2011), Renandya and Farrell (2010), Stephens (2010), and Waring (personal website, n.d.), have recently been promoting the implementation of extensive listening (EL) in the L2 classroom. This study, inspired also by the effects of extensive reading (ER), explores the effectiveness of EL by means of audio recordings of graded readers (audio books hereafter), focusing on whether L2 learners of English can improve their listening competence and acquire other linguistic knowledge, and how students perceive extensively listening to audio books.

Whilst ER is a familiar term in the language teaching repertoire, EL has received much less attention. Research into EL is nearly non-existent, and as Brown (2007, p. 15) observes, “a review of the literature on EL in a foreign language has revealed a distinct lack of statistical data.” Due to EL being a comparatively new idea, a theoretical framework is undeveloped, but it is a concept that has developed from ER (Field, 2008). Drawing on the engineering concept of bootstrapping, a process in which the results of an action are fed back to achieve greater results more quickly and with less effort, Day and Bamford (1998) have invented the term bookstrap, to emphasize that initial success in reading in a foreign language stimulates the development of a positive attitude and the growth of motivation to read more. Initial success may also lead students to read more extensively, resulting in greater achievements and pleasure in reading. The extensive reading bookstrap hypothesis could be applicable to EL, whereby learners listen to a great deal of comprehensible material, so that they are able to process the input automatically and take pleasure in listening.

Derived from ER, EL therefore refers to the access of massive amounts of easy aural input by means of television, radio, video and internet sources, or audio books and magazines (rather than listening to a particular spoken text for global understanding only), so as to improve listener automaticity in recognizing spoken texts and their listening enjoyment (Field, 2008). In a sense, it is to develop listening fluency and to foster listener autonomy.

The Effects of Extensive Reading (ER)

Theoretical Benefits of ER

Due to the lack of evidence on the effects of EL on L2 learners in the literature, reviewing theoretically and empirically relevant studies on ER is desirable to help us understand how this approach might also affect listening. Claims made by scholars
who advocate ER can be summarised as:

- Enhancing learners’ general language competence: By meeting the same patterns of letters, words, and combinations of words again and again, students process them more quickly and accurately, thus developing a sight vocabulary (Day & Bamford, 2000; Grabe, 2009).
- Increasing the learners’ exposure to the language: By reading a large number of books the “exposure gap” between L1 learners and L2 learners is reduced (Elley, 1991).
- Consolidating previously learned language (Wodinsky & Nation, 1988).
- Helping wean students away from word-by-word processing of text, and encouraging them to go for the general meaning of what they read, while ignoring details they do not fully understand (Day & Bamford, 1998; Hill, 2001).

**Empirical Evidence of the impact of ER**

There is a great amount of evidence regarding the successful impact of ER on L2 learners’ language enhancement (Elley, 1991; Elley & Mangubhai, 1983; Hafiz & Tudor, 1990; Lai, 1993a; 1993b; Lao & Krashen, 2000; Lee, 2007; Mason & Krashen, 1997; Robb & Susser, 1989; Shen, 2005; Tsang, 1996; Tudor & Hafiz, 1989). The following section briefly looks at the effect of extensive reading programs on language gain, particularly focusing on reading and vocabulary.

**Gains from Reading**

Most studies that have examined learners’ second language reading competence report gains from ER. These gains also seem to be effective in a variety of situations, ranging from underachievers (Lai, 1993a) and reluctant readers (Mason & Krashen, 1997) to students with different educational levels, and from elementary pupils in Fiji (Elley & Mangubhai, 1983) and Singapore (Elley, 1991) to secondary students in England (Hafiz & Tudor, 1990) and Hong Kong (Lai, 1993b), as well as tertiary students in Taiwan (Shen & Yuan, 2005; Lee, 2007).

In Elley and Mangubhai’s (1983) large-scale study, experimental groups made twice the expected rate of progress in reading comprehension. Another large longitudinal study in Singapore by Elley (1991) also showed significantly superior performance for REAP (Reading and English acquisition program) pupils. Shen and Yuan compared ER and guided reading, and found that after the program, the ER group had a better understanding of the main idea of an article. The same results also occurred in relation to reluctant readers in Japan, where Mason and Krashen (1997) found that an ER group scored significantly higher on understanding important ideas.

**Gains from Vocabulary**

Compared to the gains in reading skills, gains in vocabulary are less outstanding. In a study carried out in Pakistan by Hafiz and Tudor (1990), the ER participants gained significantly in their vocabulary. Similar results were found in Lao and Krashen’s (2000) study of 91 Hong Kong university freshmen, in which experimental students showed significant gains in vocabulary growth, moving one level higher than
that before the treatment, while the comparison group’s progress was very limited. In a study by Lee (2007), it was found that students who received either assigned silent sustained reading (SSR) or self-selected SSR outperformed all those who received traditional teaching on 3,000, 5,000 and 10,000 word vocabulary levels tests. A comprehensive review regarding vocabulary acquisition from ER can be seen in Horst (2005). Although each study shows some gains in vocabulary, the impact of an extensive reading program on vocabulary gain is not as encouraging as that on reading ability. This is predictable, as Nation (1997) puts the matter: “Essentially, vocabulary learning from ER is very fragile. If the small amount of learning of a word is not soon reinforced by another meeting, then that learning will be lost. It is thus critically important in an extensive reading program that learners have the opportunity to keep meeting words that they have met before” (p. 15). Above all, the ER must be done regularly and consistently for a period of time to obtain its substantial benefits.

Apart from reading and vocabulary gains from ER, other by-products from ER are improving writing (Elley, 1991; Elley & Mangubhai 1983; Lai, 1993b; Lee, 2005; Mason & Krashen, 1997; Tsang, 1996; Tudor & Hafiz, 1989), listening and speaking (Cho & Krashen, 1994; Elley, 1991; Elley & Mangubhai 1983), and attitude and motivation (Cho & Krashen, 1994; Constantino, 1994; Lao & Krashen, 2000).

Comparisons between ER and EL

Both ER and EL develop learner reading and listening skills by means of abundant written and spoken texts so that they can improve automaticity of recognition and the chunking of text, thus enabling enjoyable reading and fluent listening as early as possible. Nevertheless, some distinctions between written texts and spoken texts should be pointed out. The following is a brief summary of the differences between spoken texts (listening) and written texts (reading) taken from Chafe (1985), Raimes (1983), and Waring (personal website, n.d.):

- Spoken language is syntactically simpler than written language. The idea units in spoken language are generally joined by coordinators, such as and, but, or so, whereas written language uses more embeddings, like subordinate clauses, relative clauses, and complementary clauses.
- Spoken texts contain considerable incomplete sentences, and the idea units are shorter with an average of seven words per unit, due to the limits of short-term memory.
- Spoken language, particularly spontaneous speech, contains various disfluencies, such as fillers (you know, well, OK …), hesitations, false starts, and self-corrections, which give listeners more time to think about what has just been said and to relate it to what has gone before.
- Compared to written language, spoken language contains more colloquial expressions, slang, and nonstandard grammar, which are considered unacceptable in writing.
- Spoken language involves more face-to-face interactions, and thus tends to be more personal; whereas the writers of written language are more spatially and temporally isolated from their audiences.
- In listening, meanings may vary due to changes in areas such as stress, intonation, pitch, and volume.
- If learners cannot work out what words a string of sounds correspond to in
listening, there is nothing they can do. In reading learners can consult a dictionary.

- Listening is constrained by time. In reading learners can stop and go back easily.
- Readers can control the speed at which they read, but the listener cannot (without special recording equipment).
- In listening, sounds change from speakers to speakers because of factors such as accents and background noise, making words more difficult to acquire. In reading, the spelling/form is constant.

The comparisons between reading written texts and listening to spoken texts suggest that listening to a given text will be harder than reading the same text. It is also likely that a much larger store of receptive listening vocabulary is needed to understand a given text than when it is read because learners not only need to know that changes in areas such as stress, intonation, pitch, and volume can affect meaning but also need to know alternative pronunciations for words, e.g., res-tau-rant → res-taurant, ref-use → re-fuse. Above all, learners need to know how sounds change, elide, blend and intrude at times, so they will know what to expect.

**Comparisons between Extensive listening (EL) and Intensive Listening (IL)**

While providing L2 learners with abundant input and focusing on comprehension may be an effective way to enhance listening competence, some scholars (e.g., Field, 2008) argue that less experienced L2 listeners do not benefit from abundant input because “much of what they elicit from the input is based upon approximation or upon a principle of finding the best match” (Field, 2008, 106). Therefore, another type of listening, intensive listening (IL), has also been widely used by L2 teachers in the language classroom. According to Rost (2002), IL refers to “listening for precise sounds, words, phrases, grammatical units, specific information, and details in a style that has traditionally been used in L2 classrooms” (p. 138).

Some evidence suggests that IL practice does have positive effects upon overall listening competence, and one of the best and most common resources teachers use is “dictation” (Field, 2008), or more recently, “partial dictation” (Kuo, 2010). Dictation is usually small-scale and can be used as a remedial practice, focusing on specific listening problems, such as, linking or contraction forms, and so is a classic bottom-up approach. The findings concerning improving listening ability or other language skills through integrating IL with dictation tasks are mixed. Positive evidence can be found in Kiany and Shiramiry (2002), Kuo (2010), and Rahimi (2008); however, contradictory findings were shown in the study by Jafarpur and Yamini (1993). Overall, IL and EL are two distinct approaches; each serves different purposes in the learning process. Since IL seems to be a familiar approach whereas EL is a new term, a distinction between the two approaches is desirable. A summary of the distinction between EL and IL by is shown below (Field, 2008; Renandya, 2011, Waring, personal website, n.d.):
<table>
<thead>
<tr>
<th>Extensive listening</th>
<th>Intensive listening</th>
</tr>
</thead>
<tbody>
<tr>
<td>• listening to (or being involved in)</td>
<td>• listening for specific information</td>
</tr>
<tr>
<td>massive amounts of text</td>
<td>• listening for the exact words of a phrase or expression</td>
</tr>
<tr>
<td>• text which learners can understand reasonably smoothly</td>
<td>• listening for details</td>
</tr>
<tr>
<td>• high levels of comprehension</td>
<td>• listening to mimic a text</td>
</tr>
<tr>
<td>• listening without being constrained by pre-set questions or tasks</td>
<td></td>
</tr>
<tr>
<td>• listening at or below one’s comfortable fluent listening ability</td>
<td></td>
</tr>
</tbody>
</table>

**Language Acquisition through Listening to Stories in the L1 Context**

As previously mentioned, listening should be considered not only a comprehension skill but also an avenue to facilitate language acquisition. Research undertaken to determine the benefits of listening in a vernacular language context has been well documented in elementary pupils listening to stories read aloud by teachers. In these studies the focus has been on vocabulary acquisition. For example, two experiments carried out by Elley (1989) with New Zealand primary school pupils showed that oral story reading constituted a significant source of vocabulary acquisition. In his study, vocabulary gains by elementary students varied from 15 percent without any teacher explanation up to 40 percent with teacher explanation, with retention relatively stable in follow-up tests. Comparable results are obtained in Brett, Rothlein, and Hurley (1996), where 175 fourth graders, with teacher explanation, learned an average of three new words for each story and retention did not decay after six weeks. Another study by Senechal and Cornell (1993) even looked at 4- and 5-year-old pre-school children’s acquisition of new vocabulary from listening to a story. The results corroborated those of Elley and Bret et al.

**Language Acquisition through Listening in the L2 Context**

Unlike the consistent findings for language acquisition from listening in an L1, research into language acquisition through pure listening, i.e., listening only without any other assistance such as scripts, just as listening to one’s native language, is nearly nonexistent in an L2 context. This is because while listening, many L2 listeners usually encounter a number of difficulties (mentioned earlier). A few studies (reviewed below) seek external support (e.g., reading while listening) or adopt a “comprehension approach” by providing listeners with scripts after listening to enhance comprehension and facilitate language acquisition. However, it has to be noted that “reading while listening” is different from listening only. The former uses written texts to support and confirm listening comprehension (Chang, 2009; Field, 2008). It may be a good approach to develop auditory discrimination skills but does not help develop the strategies important for real-life listening.

Despite the dearth of research into language acquisition through “pure” listening in an L2 context, there are studies that look at language gains through different modes of input. Brown, Waring, and Donkaewbua (2008) compared learning vocabulary through reading, reading while listening, and listening only, with 35 Japanese college students studying three graded readers (The Elephant Man, One-way Ticket, and The
Witches of Pendle). Meaning-translation and multiple choice (MC) tests were used to measure their vocabulary learning rates. It was found that students consistently learned most words in the reading while listening mode, followed by reading only and then listening only. Students also found it most comfortable when the story was presented in the reading-while-listening mode, in which more students reported that the story was easy and interesting, and they knew most words and understood the story.

Similar findings were also reported in a more recent study conducted by Chang (2009) with Taiwanese college students. Chang compared L2 listeners engaged in reading while listening versus listening only in their comprehension of two short stories of equal level and length of approximately 1,500 words. Students were given an immediate post-test on story sequence for overall listening comprehension and a gap-filling test to evaluate language gains. The overall results showed that the students learning in the reading-while-listening mode achieved somewhat higher scores than with those working in the listening-only mode. Similar to the report by the Japanese students, the majority of the students in Chang’s study perceived that the reading-while-listening mode made listening tasks easier, the duration seem shorter, the stories more interesting, and they paid much better attention. The results of the above two studies imply the reading-while-listening mode could be an effective way for L2 learners to develop their listening competence and perhaps to acquire other linguistic elements as well.

Taken together, previous research into L2 listening shows that “pure listening” (listening only without any assistance) is difficult for L2 learners and to acquire linguistic knowledge through this means is even more difficult. The findings of the foregoing studies are not sufficient to determine whether L2 listeners’ listening competence can be enhanced through EL with the support of spoken scripts, in particular over a longer term. It is possible that students rely on reading and ignore listening in a reading-while-listening mode. If so, learners’ listening competence may not in fact be enhanced. To extend the scope of the studies by Brown et al. (2008) and Chang (2009), the current research intends to use well-established materials—audio books—to explore the possibility of enhancing L2 learners’ listening competence and vocabulary learning with EL as well as to investigate learners’ perceptions of it.2

The Present Study

Given the previous review of literature on language gains through listening or a combination of listening and reading, it is desirable to conduct more research in this area to shed light on the effects of EL on L2 learners’ language development. In this study, one group of participants (n = 31) adopted an EL approach, and the other group engaging in IL (n = 24) served as a control. The two groups were formed administratively according to the time slots they chose. Both groups had the class on the same day and were taught by the same teacher, who was also the researcher. The duration of each class was 60 minutes.

Both groups used audio books as class texts. The benefits of all students using the same kind of texts are as follows: 1) The teacher can guide the whole class and influence the quality of learning, and 2) It is easy to exploit content for discussion (Hill, 2001; Rosszell, 2000; Waring, 1997). Rosszell further reported that, according
to their teaching experience, the use of graded readers as class texts could stimulate students into thinking deeply about cultural, political, historical and social issues, otherwise books once finished are put back on the shelf without much further thought and quickly forgotten. Though IL is not necessarily needed for daily life, Rost (2002) argues a need to include it in listening instruction as the ability to listen intensively is an essential component of listening proficiency. They listened to the same text many times until they were familiar with the spoken forms and received a dictation test each week. Dictation was used because it has been considered an effective way of enhancing language proficiency (Jafapour & Yamini, 1993; Morris, 1983; Rahimi, 2008; Stanfield, 1985). The purpose was to explore whether EL could be as efficient as or even more effective than IL in enhancing learners’ listening competence and vocabulary acquisition in the long run.

A total of two questions are examined in this study. The first one looks at the comparative effect of different listening approaches measured by a TOEIC test (the listening section only) before and after the interventions, and the second examines whether listeners increased their general vocabulary knowledge as measured through Vocabulary Levels Tests. The specific research questions addressed are:

1. Did participants who underwent EL intervention score significantly higher in a listening test than those who underwent IL intervention?
2. Did participants who underwent EL intervention score significantly higher on vocabulary levels tests than those who underwent IL intervention?

**Methodology**

**The Participants**

The participants were 55 part-time university students from two intact classes, aged in their 20’s and 30’s. All were learning English as a foreign language in a college in Taipei, Taiwan, and were enrolled in an English assessment course in order to improve their general English competence. They were strongly encouraged, though not required, to take the Test of English for International Communication (TOEIC) at the end of the course. However, almost none were ready to face this challenge at the time the research was undertaken because their overall English competence was low, estimated at from the high beginning to intermediate level and confirmed by their TOEIC scores at pre-intervention:

**Table 1**

*Means, standard deviations, and range of listening scores for the EL and IL groups before and after the intervention*

<table>
<thead>
<tr>
<th>Time period</th>
<th>EL</th>
<th></th>
<th>IL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Range</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Pre-intervention</td>
<td>31</td>
<td>27–59</td>
<td>43.77</td>
<td>6.72</td>
</tr>
<tr>
<td>Post-intervention</td>
<td>31</td>
<td>36–69</td>
<td>50.74</td>
<td>7.45</td>
</tr>
</tbody>
</table>

*Note:* The maximum score is 100.
As shown in Table 1, the average score of the EL group was 43.77 out of 100 and that of the IL groups 42.71. There was no statistically significant demographic difference between the two groups, t(53) = .64, p = .53. None of the students needed to use English as a communication medium for their occupations. The main reasons for choosing these participants were that they had very little time to study outside the classroom and all were used to formal instruction. They attended other courses concurrently, but these had little to do with English listening or vocabulary learning. With these characteristics, any progress made would mainly come from teaching intervention.

**Study Materials and the Treatment**

All the study materials were provided by the researcher and were ranked as highly interesting by previous students who had studied them. Based on the pre-test results of the Vocabulary Levels Test, the level one graded readers of the Oxford Bookworm series were considered appropriate to start with because the participants in both groups already knew most of the 1000 and 2000 words. Fifteen audio readers were chosen for the study materials (see the Appendix). Every student had a copy of the graded reader being listened to. The treatment and other details for each group are described below.

The Extensive Listening (EL) group: In the first 14 weeks, a total of seven books were studied. As shown in the Appendix, all the books used for the first 14 weeks were very easy ones, with approximately 40 pages. As previously mentioned, these students were not used to listening to or reading English novels, the learning therefore proceeded very slowly in this first period. The researcher had to stop at the end of each chapter to make sure the students could understand the global scenarios and some important words were written on the board and explained before or after listening. However, in the second period, the remaining 12 weeks, these students moved faster than previously. They finished a total of 8 books in those 12 weeks, most of which were longer than the first seven books. To make sure they understood each story, the students were asked to work on a worksheet with comprehension questions; this was not a test but a comprehension check. Students handed in the worksheet before receiving a new book.

The Intensive Listening (IL) group: For developing their general analytical skills, the IL students received traditionally structured teaching. They studied the audio books in great detail, chapter by chapter, making sure that they understood all the utterances. During the 26-week period, they finished three stories: Jojo’s Story, Great Expectations and Amazing Grace—a rate of about one chapter per week. During each meeting, the students were given a dictation test based on the previous content studied, followed by a brief review of previous work. Dictation, considered a prototypical IL activity with emphasis on the transcription of the exact utterances (Rost, 2002), was used regularly. Students listened twice without reading the book, and then listened again while reading it. After three hearings, students checked the areas that they could not understand either by using a dictionary or by asking the teacher. If students did not raise questions, the teacher would draw their attention to the learning points in areas such as grammar, pronunciation, and culture, or asked some students to explain the scenario of the chapter. Finally, students finished the chapter by listening only, without referring to the text.
**Test Instruments and Scoring**

To investigate the gains from two different listening styles, EL and IL, listening and general vocabulary were measured by the following instruments:

- **TOEIC test**: A 100-item TOEIC test of listening comprehension was administered to all participants before and after the intervention in order to assess which group made more progress in comprehending spoken English. Since the reading section was excluded from the test, there was no score for reading comprehension. Students’ listening comprehension was reported based on the raw scores of the total number of correct items. The maximum score was 100.

- **Vocabulary Levels Tests (VLT)**: A bilingual version of VLT was adopted with a total of five levels: 1st, 2nd, 3rd, and 5th 1000 words and an academic word (AW) list. Each level contained 30 items. Due to the lack of the 4th 1000 level, the computation of the score was based on the raw scores, with one point awarded for each correct item, yielding a total of 150 points.

**Procedure**

The researcher gave the participants an orientation to the course in the first week, followed by the TOEIC listening comprehension pretest in the second week, and the vocabulary levels tests in the third week. These same measures were repeated at the end of the program. The researcher met with the students once a week, for 60 minutes each time. Between the pretest and posttest, the EL group received no formal tests, but did a practice worksheet at home after finishing each book. The IL group, on the other hand, was given a dictation test every week based on the weekly prescribed content. To capture students’ perceptions of the new learning approach, EL, students of the EL group were asked to write a report focusing on the benefits, difficulties, and disadvantages of EL at the end of the course. Since the IL group represented a traditional instructional method, these students were not asked to write a report.

**Data Analysis**

SPSS 15.0 for Windows was employed for the statistical analysis. Repeated measures of ANOVA were performed to evaluate the gains of different listening styles on student listening development, and ANCOVA was used to evaluate students’ performance on Vocabulary Levels Test because the EL group scored significantly higher statistically than the IL group. The two independent variables were Time (pre-intervention and post-intervention) and Listening Type (extensive and intensive). The written reports by the students of the EL group were translated into English to aid the interpretation of the quantitative results of the statistical analysis. The frequency of particular items reported was tallied otherwise the analysis was qualitative in nature.
Results and Discussion

RQ 1. Did participants who underwent EL intervention score significantly higher in a listening test than those who underwent IL intervention?

The descriptive statistics of pre- and post-intervention listening test scores are presented in Table 1 (repeated below), in which it is seen that the EL group scored much higher than the IL group at post-intervention. A mixed between-within subjects analysis of variance was conducted to assess the impact of the two different listening types (extensive vs. intensive) on participants’ scores on listening comprehension at pre-intervention and post-intervention. In Table 2, a significant interaction is shown between Listening Type (LT) and Time, $F(1, 53) = 6.78, p < .01$; the effect size was medium (eta squared = .08). This means both Listening Type and Time accounted for some variance for the listening scores. The main effect comparing the two types of intervention was significant $F(1, 53) = 5.77, p < .05$, eta squared =.10, suggesting a significant difference in the effectiveness of the two listening types.

Table 1
Means, standard deviations, and range of listening scores for the EL and IL groups before and after the intervention

<table>
<thead>
<tr>
<th>Time period</th>
<th>EL</th>
<th></th>
<th>IL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Range</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Pre-intervention</td>
<td>31</td>
<td>27–59</td>
<td>43.77</td>
<td>6.72</td>
</tr>
<tr>
<td>Post-intervention</td>
<td>31</td>
<td>36–69</td>
<td>50.74</td>
<td>7.45</td>
</tr>
</tbody>
</table>

Note: The maximum score is 100.

Table 2
Summary of repeated measures ANOVA for listening comprehension scores by time and listening type (LT)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within-subjects effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>538.89</td>
<td>1</td>
<td>538.89</td>
<td>21.54***</td>
<td>.27</td>
</tr>
<tr>
<td>Time × LT</td>
<td>169.73</td>
<td>1</td>
<td>169.73</td>
<td>6.78**</td>
<td>.08</td>
</tr>
<tr>
<td>Error</td>
<td>1325.96</td>
<td>53</td>
<td>25.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between-subjects effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>344.92</td>
<td>1</td>
<td>344.92</td>
<td>5.77*</td>
<td>.10</td>
</tr>
<tr>
<td>Error</td>
<td>3167.68</td>
<td>53</td>
<td>59.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p < .05; **p < .01; ***p < .001; LT: listening type

Effectiveness of Each Listening Type

Due to the interaction effect being significant, a paired-samples t-test was conducted to evaluate the effect of Time given each LT (see Table 3). There was a significant difference in listening scores between pre-intervention and post-
intervention ($M = -6.97$) for the EL group $t(30) = -5.31$, $p < .001$, indicating that students made a significant improvement after receiving abundant reading and listening input. However, the difference ($M = -1.96$) was not marked for the IL group $t(23) = -1.42$, $p > .05$.

**Table 3**
Comparisons of listening scores between pre- and post-intervention for each listening type

<table>
<thead>
<tr>
<th>Listening Type</th>
<th>$M$</th>
<th>$df$</th>
<th>$SD$</th>
<th>$SEM$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive</td>
<td>-6.97</td>
<td>30</td>
<td>7.31</td>
<td>1.31</td>
<td>-5.31***</td>
</tr>
<tr>
<td>Intensive</td>
<td>-1.96</td>
<td>23</td>
<td>6.75</td>
<td>1.38</td>
<td>-1.42</td>
</tr>
</tbody>
</table>

*Note:* The negative mean indicates that the score at post-intervention (post-test) was higher than that at pre-intervention (pre-test); **p < .01.

**Comparative Effectiveness between Listening Types**

To compare the pre- and post-intervention listening test scores of the EL and IL groups, an independent-samples t-test was further conducted. Table 4 shows that at pre-intervention, the mean difference between groups was -1.07, showing no significant difference between the two listening types $t(53) = -.64$, $p > .05$. However, at post-intervention, a statistically marked difference (MD = -6.08) was found between listening types $t(53) = -3.25$, $p < .01$, indicating that the EL group scored significantly higher than the IL group in the listening post-test.

**Table 4**
Comparisons of listening scores between listening types before and after the intervention

<table>
<thead>
<tr>
<th>Time period</th>
<th>$t$</th>
<th>$df$</th>
<th>MD</th>
<th>SE Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-intervention</td>
<td>-.64</td>
<td>53</td>
<td>-1.07</td>
<td>1.67</td>
</tr>
<tr>
<td>Post-intervention</td>
<td>-3.25**</td>
<td>53</td>
<td>-6.08</td>
<td>1.87</td>
</tr>
</tbody>
</table>

*Note:* **p < .01

That students of the EL group outperformed those of IL group at the post-intervention test was not surprising because nearly all ER research yields comparable or even better effectiveness than traditionally structured teaching (cf. Lee, 2007). Although there is no existing EL research in an L2 context that this study can be compared with, some of the written reports by students may help explain the outcomes. First, 24 out of 31 (77%) students reported that they believed their listening competence improved. This encompasses an appreciation of the advantages of being immersed in spoken language (e.g., getting used to different accents and improving pronunciation). For example, one student wrote:
In the beginning, my comprehension was low because every book was read by a different speaker, but after a few books, I got used to listening to people speaking with different accents and intonations. However, the best thing was that I could choose to listen only when the story was easy to understand, or I could read the text and listen to it at the same time when the story was difficult to understand or when I was tired. This approach is very flexible, and I really like it.4

This student’s report suggests two viewpoints. One is that EL provided listeners with the opportunity of being exposed to a variety of spoken English accents, speech rates and intonations. The other is that EL with the support of spoken scripts also allowed students to manage their own learning, monitor their comprehension and seek solutions. Apparently, reading while listening allows learners to verify the aural-written forms, which is assumed to be very helpful for low-proficiency students to develop auditory discrimination skills and for high-proficiency ones to refine word recognition skills (Osada, 2001; Mareschal, 2007; Vandergrift, 2007).

Interesting scenarios in these audio readers were reported as another major factor contributing to listening enhancement, and 18 out of 31 (58%) students reported that these stories were quite interesting, which enhanced their concentration, and that they often wished to finish a story at one time. One student reported:

Reading so many stories makes me love learning English by reading novels. I always want to know the ending of the story, so I usually did not quit until I finished a whole book. In addition to vocabulary and grammar, I also learned a lot of different cultural and moral things. Above all, I have fallen in love with these books. Also, I have never studied so many books in my English learning experience, and I will continue to read this kind of story books.

Taken together, students could monitor their comprehension during the listening process, and once comprehension broke down, they could refer to the written texts. If they could understand or wanted to challenge themselves, they listened only, without reading. By doing so, students had full control of their learning and gained a higher degree of comprehension, thereby generating high interest and concentration. Without doubt, choosing an appropriate approach plus suitable learning materials is essential to language development. Finally, the effects of EL produced similar effects as ER. In ER studies, learner reading competence is improved through reading extensively, and in this study, students’ listening comprehension was enhanced through listening extensively.

RQ 2. Did participants who underwent EL intervention score significantly higher on vocabulary levels tests than those who underwent IL intervention?

Table 5 presents the descriptive statistics of the vocabulary levels tests scores at pre-intervention and post-intervention, in which the EL group scored much higher than the IL group at pre-intervention but less so at the post-intervention. Because the 4th 1000 level was not available in the VLT, the scores were calculated based on the raw scores answered correctly by the students. Preliminary checks found that there was a significant difference between the two groups on their pre-intervention VLT scores, \(t(51) = -2.37, p < .05\), and the effect size \((d = .64)\) was moderate, so the effect of the two listening types on participants’ VLT post-test scores was assessed through ANCOVA (analysis of covariance). As shown in Table 6, there was no significant
difference between the two groups on post-intervention VLT scores, $F(1, 52) = 2.73, p = .13$, eta squared = .02. The relationship between pre-intervention and post-intervention scores on the VLT was statistically significant $F(1, 51) = 80.08, p < .005$, with a large effect size ($\eta^2 = .60$).

Table 5  
Means, standard deviations, and range for vocabulary levels tests scores for the extensive and intensive groups before and after the intervention

<table>
<thead>
<tr>
<th>Time period</th>
<th>EL</th>
<th>IL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Range</td>
</tr>
<tr>
<td>Pre-intervention</td>
<td>30</td>
<td>72–127</td>
</tr>
</tbody>
</table>

Note: The maximum score is 150.

Table 6  
Summary of the Analysis of Covariance (ANCOVA) for VLT by listening type (LT)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>337.67</td>
<td>1.00</td>
<td>337.67</td>
<td>3.58</td>
<td>.06</td>
<td>.07</td>
</tr>
<tr>
<td>Pre-VLT</td>
<td>7560.28</td>
<td>1.00</td>
<td>7560.28</td>
<td>80.08</td>
<td>.00</td>
<td>.60</td>
</tr>
<tr>
<td>LT</td>
<td>223.65</td>
<td>1.00</td>
<td>223.65</td>
<td>2.37</td>
<td>.13</td>
<td>.02</td>
</tr>
<tr>
<td>Error</td>
<td>4814.70</td>
<td>51.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>506448.00</td>
<td>54.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>12616.59</td>
<td>53.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Effectiveness of Each Listening Type**

To evaluate the effect of listening type on student scores in vocabulary levels tests at pre- and post-intervention, a paired-samples t-test was conducted. Table 7 shows that there was no significant difference between pre-intervention and post-intervention scores in the EL group, $t(30) = -.52, p > .05$, meaning that EL did not enhance the general vocabulary knowledge of students. However, there was a marked difference in the IL group, $t(22) = -3.35, p < .01$. The IL group clearly gained more in the vocabulary levels tests than the EL group.

Table 7  
Comparisons of the scores of vocabulary levels tests between pre- and post-intervention for each listening type

<table>
<thead>
<tr>
<th>Paired differences</th>
<th>M</th>
<th>df</th>
<th>SD</th>
<th>SEM</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive</td>
<td>-.94</td>
<td>30</td>
<td>9.95</td>
<td>1.79</td>
<td>-.52</td>
</tr>
<tr>
<td>Intensive</td>
<td>-6.74</td>
<td>22</td>
<td>9.66</td>
<td>2.01</td>
<td>-3.35**</td>
</tr>
</tbody>
</table>

Note: **$p < .01$**

The result of the vocabulary levels test showed that listening type did not have a significant effect on learners’ acquisition of high frequency words because they
already knew most of the 1000 and 2000 words (comprising approximately 93% of level one and level two of the graded readers in the Oxford Bookworm series according to the author’s recent study [manuscript in preparation], see also Hu & Nation, 2000), leaving only a small percentage of low frequency words to learn, a number of which are proper nouns. The IL group gained 5.75 words between the two tests but the EL group only .93 words. Although the results did not corroborate previous studies of vocabulary acquisition through ER (cf. Lee, 2007), this could simply imply that learning vocabulary through listening is limited and perhaps takes time to yield results. However, the best explanation for the outcome could be that the IL group received a dictation test every week whereas the EL group did not have any tests as such. In a recent study, Rahimi (2008), who investigated improving language proficiency by using dictation, found that his students improved the most in vocabulary rather than in listening. The second explanation for the small gain of the EL group might be that the majority of the words used in the readers were high frequency words within the 3rd 1000 level. Examining the three readers (Jojo’s story, Great Expectations, and Amazing Grace) studied by both the EL and IL groups, it seems that the three books contained higher level words; therefore, when gains in each level of words are compared, the findings showed that the IL group gained more in terms of academic words and the 3rd and 5th 1000 levels (Table 8) than other levels. Because the IL group had an opportunity to study each chapter intensively plus a dictation task every week, the higher gain of the IL group is not surprising. However, this does not mean that acquiring any vocabulary knowledge from EL is unlikely. Chang’s (2011) in-depth small-scale study with eight participants showed that her students read an average of 35 graded readers in one year and gained 8 to 22 word families (equal to 266 to 732 words). What these results imply is that EL to 15 easy audio readers may not be sufficient to bring about a significant gain in general low frequency vocabulary.

Table 8
General vocabulary gains in each level for the two groups before and after the intervention

<table>
<thead>
<tr>
<th></th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>5000</th>
<th>Academic words</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL (pre/post)</td>
<td>27.58/27.43</td>
<td>23.42/24.61</td>
<td>13.37/15.61</td>
<td>11.26/12.52</td>
<td>10.89/13.00</td>
</tr>
<tr>
<td>EL (pre/post)</td>
<td>27.67/27.94</td>
<td>24.33/24.35</td>
<td>15.60/15.16</td>
<td>13.50/13.61</td>
<td>15.90/16.39</td>
</tr>
</tbody>
</table>

Conclusion

To sum up the effects of EL versus IL on EFL learners’ listening development and vocabulary acquisition, the following was found:

- The EL group scored significantly higher than the IL group in the post-intervention listening test, implying that students’ listening competence can be enhanced through reading and listening to a number of audio books without formal instruction.
- The listening type did not show a significant effect on learners’ gains on the vocabulary levels tests between groups; however, the IL group made a
significant gain between the two time periods, while the EL group did not. The most likely reason for this might be that the three books studied by the IL group contained higher level vocabulary, and the IL group was given a dictation test every week, which helped their word building skills.

While the EL approach showed its effectiveness in enhancing EFL learners’ listening competence and vocabulary knowledge, it is important to pay attention to two factors that played an essential role in their linguistic gains and perceptions—the learning materials and the input mode. Students showed great interest in the books they studied and the interesting scenarios presented in simple language. Regarding the mode of learning, students managed their own learning by either listening only or reading while listening. If they heard incomprehensible utterances, they could always consult the written texts immediately. These two factors are keys to learning success.

Before considering the implications of this research, one characteristic of the participants’ background is worth repeating to avoid the impression that EL should be solely used in the classroom and not IL. Because the participants had been receiving much formal instruction but lacked informal exposure, the effect of EL might be due to its novelty as an intervention. Whether its effectiveness would continue in the longer term for these adult learners is unknown. However, as Nation (2007) notes, a well-balanced language learning course should involve at least four strands that require equal weighting: focused input (listening and reading), output (speaking and writing), language focused learning, and fluency development. This last strand includes all four skills, with listening being one. The listening improvement of the EL group in this study supports the position that EL can be used to develop listening fluency. Therefore, a balanced listening course should include both extensive and IL training, each serving different purposes, with the purpose of EL being to develop listening fluency.

Before ending this paper, some limitations of the study should be pointed out, along with scope for future studies. First, the study did not include reading development because the original purpose was to explore the under-researched area of listening development. Although research has well-documented reading development from ER, this study was not about “pure reading” or “pure listening”, but rather a combination of both. For L1 learners, reading while listening mainly uses listening to enhance reading ability (Beers, 1998), whereas for L2 learners, it could be the other way around or both, depending on learners’ reading and listening competence. However, listening is generally more difficult to achieve in a foreign language environment, so the majority of the students would rely on visual support (i.e., the audio script in this study) to assist their aural input. Under such circumstances, learners received dual input, and their attention was divided. Whether the effect of bimodal input increases or reduces the effect of reading or listening is unknown, and future research could include the assessment of reading development. Secondly, this study did not include students’ listening proficiency as a variable, so whether the approach has the Matthew effect—better students become better and worse students worse—is unknown. In Chang’s (2009) study, it was found that the reading-while-listening input mode did not benefit those students whose reading skills were poor because reading a written text may become a burden rather than a kind of support to aural input. Therefore, an EL approach might be more suitable for those who have acquired some reasonable amount of linguistic knowledge but still struggle with
fluency. On the contrary, lower level students might benefit more from IL because it helps learners develop bottom up listening skills, word building, matching sounds and spelling. Therefore, future research should examine the impact of different interventions on learners of different levels. Thirdly, EL in this study was limited to listening to audio recordings of graded readers, but the scope of EL should involve listening to a broader range of aural input to develop real-life listening skills. To make the research results more convincing, future research may use a broader range of materials on learners of different language proficiency levels to examine whether EL can improve their reading, listening, and vocabulary acquisition and retention.

Acknowledgements
This research was supported by a grant from the National Science Council in Taiwan (NSC97-2410-H-266-005). The author would also like to thank the anonymous reviewers for their helpful comments.

Notes
1. In this article, L2 refers to both second and foreign language.
2. The advantages of using audio recordings of graded readers can be found in Hill (2001), Rosszell (2000), and on Rob Waring’s personal website.
3. The 1st 1000 level was developed by Professor Paul Nation of Victoria University of Wellington, New Zealand. The 2nd, 3rd, 5th, 1000 words and academic words were from Schmitt, Schmitt, and Clapham (2001). The 10th 1,000 level was excluded from the test due to students’ language proficiency being low.
4. Author translation of student reports in Chinese.

References


Gains to L2 learners from extensive listening


Rost, M. (2002). *Teaching and researching listening.* London: Pearson Education.


Waring, R. (n.d.). *Starting extensive listening*. Retrieved from [http://www.robwaring.org/er/ER_info/starting_extensive_listening.htm](http://www.robwaring.org/er/ER_info/starting_extensive_listening.htm)

Watson, K., & Smeltzer, L. (1984). Barriers to listening: Comparison between students and practitioners. *Communication Research Reports, 1*, 82–87.


Appendix    Book Titles

The first 14 weeks: *Jojo’s Story*, *One-Way ticket*, *The Monkey’s Paw*, *The Elephant Man*, *Great Expectations*, *Good-bye Mr. Hollywood*, *New Yorkers*.

The second 12 weeks: *Anne of Green Gables*, *Dead Man’s Island*, *Chemical Secret*, *Amazing Grace*, *Robinson Crusoe*, *The Picture of Dorian Grey*, *The Woman in Black*, *The Hound of the Baskervilles*

(NB.: The underlined titles were studied by both the EL and IL groups.)