

USE IT OR LOSE IT; L2 LISTENING ATTRITION OVER A SEMESTER

Associate Professor Dr. Omar Karlin
Toyo University.
email: o_karlin@yahoo.com

Adjunct Professor Dr. Sayaka Karlin
Toyo Gakuen University.
email: sayakakarlin@yahoo.co.jp

ABSTRACT

Forty Japanese university students in an intercultural communication program at a private university in Tokyo were split into two groups to determine the L2 listening attrition that occurred over a semester. The treatment group was given intensive listening homework which involved transcribing a popular American television show, while the control group was not given any listening homework. After calculating Rasch person measures for each student, and conducting paired-sample t-tests, results indicated a significant decrease in L2 listening ability relative to other students for the control group. Of the five variables being measured on an end-of-semester listening exam; *all items*, *all listening items*, *non-textbook listening items*, *textbook listening items*, and *vocabulary items*, a significant decrease was evident in the variables *all items* ($t(20) = 3.61, p < .05$), *all listening items* ($t(20) = 2.89, p < .05$), and *vocabulary items* ($t(20) = 2.90, p < .05$) for the control group. No significant decreases in ability were evident for the treatment group.

Keywords: EFL, listening, homework, transcription, L2 attrition

INTRODUCTION

The phrase *use it or lose it* is often used but sometimes not fully understood, mostly because of the ambiguous *it* pronoun. What is *it*? *It* can be used to cover a large number of situations, with varying degrees of truthfulness to the phrase. For anyone that has made a resolution to exercise, *it* can refer to the process of staying fit, building muscle, and/or losing weight, and usually the phrase *use it or lose it* has never rung truer. Stopping an exercise routine can result in a reversion to the pre-exercise self within weeks or months (i.e. not being fit, lost muscle, and the gaining of weight). However, there are instances when the phrase *use it or lose it* rings hollow, such as with riding a bicycle. People can go long periods of their life without ever riding a bicycle, yet rarely revert to their pre-bicycle-riding selves. In fact, the well-known phrase *it's like riding a bike* basically means the opposite of *use it or lose it*; even if you don't use it, you won't lose it. Many activities, including using a second language, fall somewhere between these two poles.

Among the four skills, listening is the skill that many teachers have the most difficulty with teaching. With productive skills such as writing and speaking, learners produce something that teachers can assess, and with reading, learners at least have a reading passage which is physically present, which the teacher can reference during a lesson. However, with listening, teachers must impart a skill in which the students produce nothing, and there is no physical remnant of the listening passage to reference. Faced with these challenges, many teachers resort to *assessing* listening, rather than *teaching* listening. The chief teaching methodology than many listening teachers employ is to ask comprehension questions to their students, which does little to improve ability. If teachers do not take a more active role in trying to improve L2 listening ability through transcription and reflection (or other activities), L2 attrition can occur. Even though receptive skills are thought to be relatively stable, this study demonstrates that L2 listening attrition can occur within a single semester if students are not given adequate listening homework.

Literature Review

entry into university

The zenith of studying intensity for Japanese students is high school. From the time that students enter school at five years old, there is a gradual intensification of competition to matriculate into well-regarded schools (Mawer, 2015; Mori & Baker, 2010). This intensification modestly begins in elementary school, however because the choice of public junior high school for many is restricted to geographical proximity, competition is more modest. However, as students enter junior high school, competition intensifies significantly as there are no restrictions for high school matriculation, meaning that students can enrol at any high school that will have them. Junior high school students study intensely in order to secure admission into a prestigious high school. In fact, the importance

of success on high school entrance exams has resulted in a large system of shadow education in Japan (an industry of 93.6 billion yen in 2013; Mawer, 2015), in which parents pay for their children to attend private evening classes to supplement their normal schooling. Once students matriculate into a high school, those that plan on continuing onto university must successfully navigate one final period of entrance examinations, those for university matriculation. These exams represent the final entrance exams that most students will take, and doing well and being admitted into a good university can ensure their future professional and social success. As a result, some students who do poorly will actually decline an offer to enrol in a lesser university so they can study intensely for a full year (without being enrolled in any school) and attempt the entrance exams again (Koyama et al., 2014). Once students have finished entrance exams and been admitted into a university, their motivation to study often declines significantly (Berwick & Ross, 1989). To use an analogy, studying for the high school entrance exams is like a roller coaster that has reached the top of a precipice, once those exams are finished, the motivation and effort for many will plunge.

L2 language attrition

If the studying intensity of students peaks in the final year of high school, with motivation and effort plunging once they enter university, a reasonable question would be to wonder if their L2 abilities will adopt a new downward trajectory. The increase of second language ability is referred to as second language acquisition, while the opposite, the decrease of second language ability is referred to as second language attrition. Just as with second language acquisition, second language attrition is highly dependent on contextual factors, yet there are some central hypotheses that frame much of the existing research in this area of inquiry. One of these is the regression hypothesis, which suggests that the first skills learned are the last to be lost (Wei, 2014), thus, because most people learn to read before they speak, attrition tends to occur with productive skills before receptive skills (Hanson, 1999; Wei, 2014). Some research has indicated that receptive skills may even increase while productive skills have started to atrophy (Edwards, 1976; Welton, 1987). An additional hypothesis that suggests productive attrition would occur before reading attrition is the retrieval fail hypothesis, which suggests that forgetting a language can stem from the lack of visual stimuli (Wei, 2014). With reading, the letters, words, and paragraphs are omnipresent and can assist the reader in retrieving language, however, with the productive skills, language is generated by the learner; thus there is no pre-existing visual stimuli from which to rely on.

Tomiyama (1999) looks at L2 attrition within a Japanese context, and comes to several relevant conclusions. Chiefly among Tomiyama's conclusions is the assertion that L2 attrition, while generally similar, is not uniform, and can occur at different speeds, and target different areas of attrition, depending on the individual. The primary subject of her study, a Japanese elementary student named Ken, exhibited attrition first in the areas of fluency and lexical retrieval, while receptive comprehension remained strong throughout the 19-month period of the study. L2 attrition first started to appear at six months, which was consistent with existing research. Tomiyama notes that two other subjects in her study had L2 atrophy at a different pace, and interestingly, the two boys in her study developed a fluency atrophy compensation strategy of codeswitching, while the girl in her study developed a fluency atrophy compensation strategy of hesitations and pauses. However, it should be noted that the subjects in Tomiyama's study were relatively young, and returnees. Research on second language acquisition suggests that children are much more capable of learning a language than adults (Bialystock & Hakuta, 1999), which might suggest the opposite for attrition. Because adults struggle more when learning an L2, attrition might be more pronounced for them, as opposed to children who might possess a more stable understanding of an L2 because of their relative mastery during acquisition.

In another Japan-based study, Yoshitomi (1999) suggests less identifiable L2 attrition than Tomiyama (1999), however, she does provide explanations for this. Yoshitomi suggests that the four subjects of her study, also of elementary school age, did not exhibit significant L2 attrition in any of the subskills examined in her study, such as understanding of verb morphology, articles, and lexicon, after approximately one year. However, Yoshitomi contends that while the observable L2 attrition is light, there are two major caveats with this finding. The first is that even though L2 attrition may be minor in some linguistic subskills, the cumulative effect on fluency can be significant, both in terms of lacking linguistic skills and lacking confidence to engage in conversations for practice. Further, Yoshitomi suggests that assessment may not be accurately capturing L2 attrition in that many L2 learners develop compensation strategies in the face of their attrition, and these compensation strategies are not recognized as attrition. Yoshitomi cites examples such as an L2 learner who retrieves lexicon that might not be perfectly suitable for a situation, because they can no longer remember the most suitable lexicon, or a reticent speaker who is deemed a little shy rather than somebody who is refusing to participate because they recognize their own attrition. In these instances, attrition is masked through compensation strategies. Finally, Yoshitomi also suggests that the commonly agreed upon principle that productive skills atrophy before receptive skills may also be a victim of poor assessment.

Typically, the attrition of productive skills is due to a collective attrition in smaller linguistic sub-skills, and are easier to identify with assessment than receptive skills. However, Yoshitomi suggests that language skills are not isolated, and that productive and receptive skills are interrelated, so attrition in productive skills likely means a similar attrition in receptive skills. The poor assessment of receptive skills is reflected in the generally poor state of teaching listening in most language classrooms.

Teaching listening

Of the four language skills of reading, writing, listening, and speaking, listening involves the least explicit teaching. With reading, learners are taught to recognize familiar prefixes and suffixes, how to infer meaning from root words, and how to skim and scan text. With writing, learners are taught how to construct sentences, paragraphs, and essays. With speaking, learners practice using situational phrases for agreeing, disagreeing, clarifying, and other conversational necessities. However, with listening, more often than not, learners listen to a passage and are asked comprehension questions on what they just heard (Brown, 2011; Field, 2008; Goh, 2008; Siegel, 2014; Thorn, 2009). The focus is almost exclusively on the product of ability, rather than the process of learning. As Schmidt (2016) suggests, asking comprehension questions in a listening class is not the *teaching* of listening but rather the *assessment* of listening.

Rost (2014) suggests that there are three main components to listening, the affective domain, the cognitive domain, and the interpersonal domain. The affective domain largely deals with the willingness of the learner to make a connection with the L2 target culture, the cognitive domain largely deals with the learner's ability to recognize lapses in their listening skills and to not become derailed while actively engaged in listening, and the interpersonal domain largely deals with the learner's ability to navigate a personal interaction that requires them to listen. When viewing these three domains of listening through the context of a Japanese learner, it is important to note that these domains are not present in equal measure. In Japan, where there are few English-speaking students or residents (compared to other developed countries), the interactional domain has limited scope because there are simply fewer opportunities to interact. As a result, the affective and cognitive domains take on greater importance for L2 listening acquisition. With regard to the affective domain, the internet and the abundance of available English-speaking media have given teachers more opportunities for exposing students to the target culture and creating meaningful cultural connections. In fact, the use of authentic listening sources can have a motivating effect on learners (Mousavi & Irvani, 2012; Thorn, 2009). With regard to the cognitive domain, learners who can relearn auditory properties, and retrain themselves in strategies that compensate, approximate, reframe, skip, and substitute elements of what they have heard, should be able to improve their listening ability (Rost, 2014).

However, integrating these three domains of listening into a coherent and effective classroom strategy eludes many listening teachers. The cognitive domain, in particular, should be targeted by teachers. While it may be practically impossible to facilitate the interpersonal domain for Japanese students due to the lack of English-speaking interaction opportunities in Japan, and facilitating the affective domain requires establishing a meaningful connection with the target culture, which can be difficult to do given the grammar-focus of Japan's English classrooms, the cognitive focus would seem to lend itself to the bottom-up nature of grammar-based teaching that is omnipresent in Japan. Unfortunately, most listening practice, not only in Japan but around the world, is done with a top-down approach, focusing on comprehension rather than words and sounds (Ngo, 2016; Thorn, 2009). Bottom-up processing is an important element to improved listening ability (Field, 2003; Kiany & Shiramizu, 2007; Renandya & Farrell, 2011). Listening journals in which students focus on identifying the units of sentences, and reflect on their learning, can be an effective way of utilizing a bottom-up approach to teaching listening (Schmidt, 2016). Similarly, dictation can be modified in many ways to be more interesting (Mumford, 2016), yet neither approach is commonplace in language classrooms. While many teachers acknowledge the importance of listening and have a desire to employ teaching strategies that are informed by research and help students improve the process of listening, few are able to put into practice these beliefs and instead rely on the use of textbooks and assessment in listening classes (Graham, Santos, & Francis-Brophy, 2014, Karimi & Nazari, 2017; Ngo, 2016).

For those teachers that would like to construct listening activities that focus on the process of listening, rather than the product of listening, Chang, Wu, and Pang (2013) identify several of the main difficulties that learners face while listening. In their meta-analysis, as well as in their own research, Chang, Wu, and Pang report several difficulties faced by learners; speech rate, pronunciation, text type (such as a lecture or a conversation), hesitations, task type (such as local or global questions, multiple-choice questions, or comprehension questions), contextual support, background knowledge, and other factors. Their recommendations for tasks are that they align with learners interests in order to motivate them, that they offer contextual support (such as opportunities to repeat what they heard or to see visual images), that students are exposed to different types of listening texts (such as monologues and conversations, formal and informal), that students hear different types of accents and speech rates,

and that comprehension testing be minimized in order to remove anxiety from their listening experience (Chang, Wu, & Pang, 2013). In another study examining listening tasks, Hosogoshi (2016) demonstrated that when learners were given English captions to videos, there was no significant difference in comprehension with those that were not given captions. However, research by Chang and Millett (2014) suggests that reading a text while simultaneously listening to an audio recording of the text can help with listening comprehension, although it should be noted that the students in this research were beginner-level, which may account for the discrepancy with Hosogoshi's (2016) findings.

Research Question

1. To what extent is student performance on an end-of-semester listening exam dependent upon the amount of authentic listening homework that they receive during a semester?

Participants

Participants were drawn from two English classes of equal proficiency, as determined by a TOEFL ITP placement test. Both classes were categorized as high-advanced, which meant that students had a TOEFL iBT score in the range of 65-80, with several students living abroad before enrolling at university. Students with experience living abroad, either short-term or long-term, represented approximately a fifth of all students in these classes, and experiences abroad occurred in a variety of countries, both English-speaking and non English-speaking. The two classes included 19 and 21 students, respectively. All 40 students involved in this study were in their first year at a major private university in central Tokyo. Students were majoring in intercultural communications and had eight class periods of English study, comprising four courses, per week. The students' four courses included a listening course, a speaking course, a reading & writing course, and one elective (usually a TOEFL or intercultural comparison course). Courses lasted a full year, with students receiving a spring semester grade and an autumn semester grade.

An additional 75 students also took the final exam, however, these students were in low-advanced classes, their TOEFL iBT score was in the range of 55-65, and they took an exam that was significantly different than the exam that the high-advanced students took. The high-advanced and low-advanced exams had 56 questions each with 26 shared questions. As a result, the results of the low-advanced exam were grouped with the results of the high-advanced exam in order to conduct a Rasch analysis (to obtain Rasch person measures across the entire advanced level). However, the low-advanced students were not included in this study.

Instrument

Data was collected at the end of the students' first semester, and again at the end of their second semester, via the final exam held for students at the end of each semester. Each exam consisted of 56 vocabulary and comprehension questions. Of the 56 questions, 14 were vocabulary questions which required students to match a word with the appropriate definition, 12 were comprehension questions based on two lectures from the teacher's version of the class textbook, and 30 were comprehension questions based on five lectures from online sources. All of the listening comprehension questions were multiple-choice questions with four distractors. Each exam took one hour to complete.

Person ability measures derived from the Rasch measurement model were calculated from the raw scores of each student on each exam. The first semester exam served as the pretest for this study, while the second semester exam served as the posttest. The resulting difference between the two exams in five variables (all items, listening items, non-textbook listening items, textbook-listening items, and vocabulary items) represented the effect of the treatment in the treatment group and the control group.

METHOD

Each class was taught by a native speaker of English who used the same textbook and followed the same weekly schedule of textbook progression. However, the homework given by each teacher differed. For the control group, the teacher assigned no listening homework and did all listening practice during class time. For the treatment group, the teacher assigned six "listening logs" and one mid-term listening assignment.

Each listening log required students to watch an episode of the CBS reality television show *The Amazing Race*. *The Amazing Race* is a television series that, as of this writing, has broadcast 29 seasons and been on CBS since 2001. Each year of broadcast includes two seasons, and each season is comprised of 12 or 13 episodes. Over the course of a season, 11 or 12 teams of two (usually close friends or family members) race to different locations around the world. Each episode involves the distribution of clues, the race to a new location, and the elimination of the last-place team from the competition. The one remaining team at the end of a season wins \$1,000,000. While

students watched an episode of *The Amazing Race*, they would transcribe one A4 page of dialogue from the episode, as well as write one A4 page of reflection, consisting of four questions. These reflection questions included giving a summary of the episode, explaining the details of their listening (such as the location of their listening, duration of their listening, their comprehension, and their impression of the listening experience), listing five new vocabulary and three new phrases from the episode, and assessing their listening improvement in the course. Students usually took between two and three hours to complete a single listening log, with the first listening log of the first semester taking some students up to ten hours to complete, and the last listening log of the second semester taking most students 90 minutes to complete. Each of the six listening logs were assigned at intervals of two or three weeks, and students would have one week to complete each listening log. This study's use of *The Amazing Race* as the basis for a transcription activity was thought to accomplish many of Chang, Wu, and Pang's (2013) suggestions that stemmed from their research and their meta-analysis. *The Amazing Race* is a travel competition program and many of the students in this study were intercultural communications majors, which naturally oriented them towards travel and other cultures, watching episodes for homework provided the contextual support of repeat viewing opportunities and visual images to assist with comprehension, each episode offered several types of listening texts such as the narrator's monologue summary of the previous episode and the conversational banter between contestants, the variety of contestants involved in the program offered different accents and speech rates, and the activity did not involve testing so as to minimize the anxiety usually associated with comprehension testing. Also, using Hosogoshi's (2016) findings, it was decided that not providing students with English captions would be advantageous, since captions provided no significant increase in understanding, and that the presence of English text at the bottom of the video image could prevent students from fully relying upon their listening ability.

The mid-term listening assignment followed the same format as the listening logs, except students had to listen to an academic lecture (of their choice) rather than an episode of *The Amazing Race*. Students had to choose an academic lecture from TED.com, a website that hosts thousands of lectures of varying length on academic, scientific, and cultural topics, and complete a one-page transcription and answer four reflective questions, similar to their listening log homework.

Over the course of their second semester, the control group had no intensive listening homework assignments, while the treatment group had seven intensive listening homework assignments (six listening logs and a mid-term assignment).

RESULTS

Paired-sample t-tests were conducted to determine L2 listening proficiency changes after a semester of studying, as determined by a final exam, after assigning intensive listening homework (treatment group) and no listening homework (control group). The descriptive statistics for the treatment group's pretest and posttest Rasch person measure scores are shown in Table 1.

Table 1. *Rasch Person Measures for Treatment Group, Descriptive Statistics and Correlations in Exam Variables*
Note: $n=19$; $*p < 0.05$

	Spring semester			Autumn semester			Corr.	p
	M	SD	Std. Error M	M	SD	Std. Error M		
All items	1.10	.74	.17	.90	.64	.15	.80	.00*
All listening items	.64	.75	.17	.46	.65	.15	.81	.00*
Non-textbook listening items	.13	.71	.16	.14	.69	.16	.81	.00*
Textbook listening items	1.09	1.14	.26	1.23	.94	.22	.45	.05
Vocabulary items	3.61	1.13	.26	4.06	1.25	.29	.39	.10

Significant correlations were observed in three variables, *all items* ($r(18) = .80, p < .05$), *all listening items* ($r(18) = .81, p < .05$), and *non-textbook listening items* ($r(18) = .81, p < .05$). The results of the paired-sample t-test are shown in Table 2.

Table 2. Paired-sample t-tests of Rasch Person Measures for Treatment Group, Mean Difference and Significance in Exam Variables

	Paired differences					T(18)	P	H
	M diff.	SD	Std. Error M	95% interval Lower	confidence Upper			
All items	1.10	.74	.10	-.02	.41	1.92	.07	.43
All listening items	.64	.75	.10	-.03	.39	1.78	.09	.21
Non-textbook listening items	-.01	.43	.10	-.22	.20	-.11	.91	-.02
Textbook listening items	1.09	1.14	.25	-.67	.39	-.55	.59	.53
Vocabulary items	3.61	1.13	.30	-1.08	.18	-1.49	.15	1.23

Note: $n=19$; $*p < 0.05$; Morris (2008) used in calculation of effect size (H).

No significant differences were observed between the treatment group's pretest Rasch person measure scores and the treatment group's posttest Rasch person measure scores. The descriptive statistics for the control group's pretest and posttest Rasch person measure scores are shown in Table 3.

Table 3. Rasch Person Measures for Control Group, Descriptive Statistics and Correlations in Exam Variables

	Spring semester			Autumn semester			Corr.	p
	M	SD	Std. Error M	M	SD	Std. Error M		
All items	1.02	.50	.11	.55	.55	.12	.37	.10
All listening items	.60	.52	.11	.29	.58	.13	.60	.00*
Non-textbook listening items	.04	.55	.12	.06	.58	.13	.52	.02*
Textbook listening items	1.27	1.15	.25	.80	.91	.20	.39	.08
Vocabulary items	3.29	1.36	.30	2.18	1.41	.31	.21	.37

Note: $n=21$; $*p < 0.05$

Significant correlations were observed in two variables, *all listening items* ($r(20) = .60, p < .05$), and *non-textbook listening items* ($r(22) = .52, p < .05$). The results of the paired-sample t-test are shown in Table 4.

Table 4. Paired-samples t-tests of Rasch Person Measures for Control Group, Mean Difference and Significance in Exam Variables

	Paired differences					T(20)	P	H
	M diff.	SD	Std. Error M	95% interval Lower	confidence Upper			
All items	.47	.59	.13	.20	.74	3.61	.00*	.43
All listening items	.31	.49	.11	.09	.54	2.89	.01*	.21
Non-textbook listening items	-.01	.55	.12	-.26	.24	-.11	.92	-.02
Textbook listening items	.47	1.16	.25	-.06	1.00	1.86	.08	.53
Vocabulary items	1.10	1.74	.38	.31	1.90	2.90	.01*	1.23

Note: $n=21$; $*p < 0.05$; Morris (2008) used in calculation of effect size (H).

Three significant differences were observed with the control group, indicating that not giving listening homework during a semester resulted in the loss of L2 listening ability relative to other students in the department. The variables that exhibited a significant decrease between the pretest and posttest scores included *all items* ($t(20) = 3.61, p < .05$), *all listening items* ($t(20) = 2.89, p < .05$), and *vocabulary items* ($t(20) = 2.90, p < .05$).

DISCUSSION

While the results of the paired-sample t-tests indicated that the control group suffered from L2 listening ability attrition, it must be noted that the significant decrease in the control group's posttest exam scores were not raw scores. When conducting this analysis, the exam's raw scores were converted to Rasch person measures which are a measure of relative ability among all of the exam-takers. Also, if the posttest was more difficult than the

pretest, it is possible that the decreasing person measure does not represent L2 attrition. The control group may have maintained their L2 listening ability and just been faced with a more difficult test, resulting in the decreased Rasch person measure. However, the Rasch person measure does indicate that relative to the other students who took the exam, the ability of the control group decreased. This explanation would be consistent with previous findings that suggest that L2 language attrition tends to not occur as strongly with receptive skills such as reading and listening. Nevertheless, the Rasch person measure of the L2 listening ability of the control group did decline significantly in relation to their peers in the entire advanced-level. While the treatment group's L2 listening ability also declined relative to their peers in the entire advanced-level, it was not significant. While the L2 listening ability of both high-advanced classes decreased, only the control group's L2 listening ability decreased to a significant degree. To account for the decrease in relative ability by both groups, the L2 listening ability of the low-advanced students increased, likely as a result of the intensive nature of the L2 program at this university, and the tendency for advanced classes to use the same materials for high-advanced and low-advanced students.

It does appear that assigning intensive listening homework had a positive effect on the treatment group. The results of this study only offer a limited picture of how the treatment group may have benefited from the treatment, and it is likely that the intensive listening homework also had ancillary benefits on their other English skills, such as reading, writing, and speaking. By watching *The Amazing Race* and improving their recognition of native-speaker speech patterns (such as speech, rate, idiomatic language, pauses and hesitations, and accents), the treatment group may have been better able to initiate conversations with the international students in their program, international residents in their community, or engage in study-abroad and travel opportunities, which would likely further improve their speaking, and possibly their reading and writing. Without another measure to confirm this theory, it is difficult to say with certainty, but it seems likely that the TOEFL iBT scores for the treatment group would have improved across all language skills. Just through observation, it seemed that most of the students in the treatment group, more than the control group, were accepted into study-abroad programs in their second-year at university. This finding, however, has not been substantiated with data from the university. Connected to the possibility that intensive listening led to ancillary improvement in other skill areas, it is likely that the treatment group enjoyed affective benefits from watching *The Amazing Race*. As was mentioned earlier, students in this intercultural communication program tended to be interculturally-oriented, with a strong interest in meeting people from other cultures and travelling to different countries. The tendency of *The Amazing Race* to present global locales in an appealing way, as a travel documentary might, as well as creating a compelling narrative with the native-speaking contestants, probably increased the desire of many students in the treatment group to go abroad, travel to some of the countries featured in *The Amazing Race*, and meet like-minded native-speakers. Recalling Rost's (2014) domains of listening, the affective domain, in which L2 speakers create meaningful connections with the L1's culture, is one of the most crucial aspects to developing listening ability. After reading the responses to the reflection questions for each listening log assignment, it was clear that students were really attracted to many of the travel destinations depicted in the program. Many students indicated that they wanted to travel to the places on the show. Additionally, many students, if not all, had clear rooting interests for the contestants on the show, cheering for some to win and for others to be eliminated. It is rare for textbooks to elicit this kind of emotion, so it seems that the treatment group likely enjoyed affective benefits that the control group did not. Perhaps future research could add an affective survey to determine the extent of affective benefits that come with this type of intensive listening homework.

It was interesting that several students in the treatment group had a strong desire to see transcripts of the episodes that they had watched. One of the goals of this activity, besides the development of bottom-up processing skills and the establishment of meaningful connections with the target culture, was to simulate the experience of real-time conversations, that is, to listen and not be entirely certain of what is being said. Obviously when somebody is engaged in a conversation, they do not have the option of stopping the conversation and looking at a transcript to confirm understanding. It was thought that not having access to a transcript would develop the students' tolerance for ambiguity, and that increased tolerance of ambiguity could lead to linguistic and affective gains (Dewaele & Wei, 2013).

On a purely qualitative basis, using *The Amazing Race* as the basis for an intensive listening activity was successful. Students spent a significant amount of time practicing their listening and completing the transcription and reflection portions of the assignment, yet they never seemed to burn out from the activity. From this researcher's experience, teachers need to be cognizant of how much homework they can assign to students before they become overwhelmed or demotivated. For most students, the first listening log took between 5-10 hours to complete, however, the effort of students did not wane in subsequent listening logs. The length and detail of transcripts was constant throughout the course, and responses to reflection questions remained thoughtful through all seven listening assignments in the second semester (as well as the seven listening assignments in the first semester).

It seems that the cultural closeness of this activity and the narrative constructed within the show had a powerfully motivating effect on students, which was consistent with research that suggested authentic materials can provide motivational benefits, and should be used whenever possible (Mousavi & Irvani, 2012; Ngo, 2016; Thorn, 2009). It seems unlikely that using materials from the textbook as the basis for intensive listening activities would have held the same sway over students for an entire semester (or two semesters!), although this would need to be confirmed with an affective survey. With the internet making music, magazines, movies, television, newspapers, and other cultural artifacts readily available to teachers, modifying L2 listening activities to include authentic materials should be a guiding tenet of listening teachers and textbook publishers.

Limitations

While the results of this study offer an insight into how L2 listening attrition can occur over just a single semester, there are several limitations to this study that readers must be aware of. First, the characterization of L2 listening attrition in this study is relative not absolute. While the exam scores for the control group decreased significantly, this does not necessarily mean that absolute L2 listening ability decreased. If the pretest was easy and the posttest was difficult, it is possible that absolute L2 listening ability actually increased for the control group but is hidden by the discrepancy in the difficulty levels of the two exams. The significant decreases in L2 listening ability observed in this study are relative to other students within the advanced level of the program. The Rasch mathematical model calculates a person ability measure by incorporating the performance of all of the exam's respondents into a predictive model, so relative to other students in the advanced level, the control group's ability declined, yet this does not necessarily mean in an absolute sense. Second, as mentioned in the discussion, the listening exam used as the pretest and posttest of this study only gives a limited view of L2 proficiency. The intensive bottom-up processing involved with transcriptions likely gave the treatment group additional L2 development benefits, such as greater confidence in listening to native speakers which might have led to increased conversation practice (and improved speaking ability) or incidental vocabulary acquisition (and improved reading ability). Utilizing additional sources of proficiency data, such as TOEFL scores, would give a more complete picture of L2 acquisition stemming from intensive listening homework. Third, the scope of this research study, involving only 40 students over a single semester, was limited. Statistical results would be more reliable with more participants. Moreover, most research has tended to show that L2 attrition generally does not become apparent until after six months, thus a longer timeline for this study might have revealed more pronounced decreases in L2 listening attrition.

CONCLUSION

The adage *use it or lose it* holds true for many aspects of life, from maintaining physical fitness to cognitive acuity. It has been suggested that receptive skills are more insulated for L2 attrition than productive skills, however this study demonstrated that L2 listening attrition, in a relative sense, can occur after just a single semester. It is essential for teachers to prevent this by adopting more progressive strategies for teaching listening. Simply relying on in-class checks of comprehension, and not utilizing outside authentic materials for homework, can have a detrimental effect on students. Luckily, with the increasing connectivity enjoyed around the world, there has never been a greater abundance of authentic materials available to teachers to incorporate into their lessons and to help their students build their L2 skills.

REFERENCES

- Berwick, R. & Ross, S. (1989). Motivation after matriculation: Are Japanese learners of English still alive after exam hell? *JALT Journal*, 11(2), 193-210.
- Bialystock, E. & Hakuta, K. (1999). Confounded Age: Linguistic and Cognitive Factors in Age Differences for Second Language Acquisition. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Lawrence Erlbaum Associates: Mahwah, New Jersey.
- Brown, S. (2011). *Listening myths: Applying second language research to classroom teaching*. Ann Arbor: University of Michigan Press.
- Chang, C-S. & Millett, S. (2014). The effect of extensive listening on developing L2 listening fluency: Some hard evidence. *ELT Journal Volume*, 68(1), 31-40. doi:10.1093/elt/cct052
- Chang, C-S., Wu, W-P., & Pang, C-L. (2013). Second language listening difficulties perceived by low-level learners. *Perceptual & Motor Skills: Learning & Memory*, 116(2), 415-434. doi: [10.2466/22.23.PMS.116.2.415-434](https://doi.org/10.2466/22.23.PMS.116.2.415-434)
- Dewaele, J.-M., & Wei, L. (2013). Is multilingualism linked a higher tolerance of ambiguity? *Bilingualism: Language and Cognition*, 16, 231-240. doi:10.1017/s1366728912000570

- Edwards, G. (1976). Second language retention in the Canadian public service. *The Canadian Modern Language Review*, 32, 305-308.
- Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, 57(4), 325–334. doi:10.1093/elt/57.4.325
- Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.
- Goh, C. (2008). Metacognitive instruction for second language listening development: theory, practice and research implications. *RELC Journal*, 39, 188-213. doi:http://dx.doi.org/10.1177/0033688208092184.
- Graham, S., Santos, D., & Francis-Brophy, E., (2014). *Teacher beliefs about listening in a foreign language. Teaching and Teacher Education*, 40, 44-60. doi: https://doi.org/10.1016/j.tate.2014.01.007
- Hosogoshi, K. (2016). Effects of captions and subtitles on the listening process: Insights from EFL learners' listening strategies. *JALT CALL Journal*, 12(3), 153–178
- Karimi, M. & Nazari, M. (2017). The Congruity/Incongruity of EFL Teachers' Beliefs about Listening Instruction and their Listening Instructional Practices. *Australian Journal of Teacher Education*, 42(2). doi:http://dx.doi.org/10.14221/ajte.2017v42n2.5
- Kiany, G. R. & Shiramiry, E. (2002). The effect of frequent dictation on the listening comprehension ability of elementary EFL learners. *TESL Canada Journal* 20(1): 57–63. doi: https://doi.org/10.18806/tesl.v20i1.938
- Koyama, A., Matsushita, M., Ushijima, H., Jono, T., & Ikeda, M. (2014). Association between depression, examination-related stressors, and sense of coherence: The ronin-sei study. *Psychiatry and Clinical Neurosciences*, 68, 441–447. doi:10.1111/pcn.12146
- Mawer, K. (2015). Casting new light on shadow education: Snapshots of juku variety. *Contemporary Japan* 2015, 27(2), 131–148. doi:10.1515/cj-2015-0008
- Mori, I. & Baker, D. (2010). The origin of universal shadow education: what the supplemental education phenomenon tells us about the postmodern institution of education. *Asia Pacific Educational Review*, 11, 36-48. doi:10.1007/s12564-009-9057-5
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2), 364-386. http://doi.org/10.1177/1094428106291059
- Mousavi, S. A. & Irvani, H. (2012). The effect of authentic versus non-authentic aural materials on EFL learners' listening comprehension. *English Language and Literature Studies*, 2(1), 21–27. doi:10.5539/ells.v2n1p21
- Mumford, S. (2016). *Intensive listening; Some possibilities*. Retrieved from www.modernenglishteacher.com
- Ngo, N. T. H., (2016). The Impact of Listening Strategy Instruction on Listening Comprehension: A Study in an English as a Foreign Language Context. *Electronic Journal of Foreign Language Teaching*, 13(2), 245–259.
- Renandya, W. A. & Farrell, T. S. C. (2011). “Teacher, the tape is too fast!” Extensive listening in ELT. *ELT Journal* 65(1), 52–59.
- Schmidt, A. (2016). Listening journals for extensive and intensive listening practice. English Teaching Forum.
- Siegel, J. (2014). Exploring L2 listening instruction: examinations of practice. *ELT Journal*, 68(1), 22–30. doi:0.1093/elt/cct058
- Thorn, S. (2009). Mining listening texts. *Modern English Teacher* 18(2): 5–13.
- Tomiya, M. (1999). The First Stage of Second Language Attrition: A Case Study of a Japanese Returnee. In L. Hansen (Ed.), *Second Language Attrition in Japanese Contexts*. New York: Oxford University Press.
- Wei, J. (2014). Selectivity of second language attrition. *Theory and Practice in Language Studies*, 4(8), 1603-1608. doi:10.4304/tpls.4.8.1603-1608
- Weltens, B. (1987). The attrition of foreign-language skills: a literature review. *Applied Linguistics*, 8, 22-36.
- Yoshitomi, A. (1999). On the Loss of English as a Second Language by Japanese Returnee Children. In L. Hansen (Ed.), *Second Language Attrition in Japanese Contexts*. New York: Oxford University Press.