

Article

The Effects of Transcript Use on Advanced ESL Listening Comprehension

K. James Hartshorn*

Brigham Young University, USA

Chirstin Stephens

Southern Utah University, USA

Received: 4 May 2023; Accepted: 28 June 2023; Published: 5 November 2023

Abstract

This study examined the efficacy of transcript use with repeated listening exercises to improve advanced ESL student listening comprehension. Thirty-one students enrolled in an intensive English program from various L1 backgrounds participated with 15 in the treatment group and 16 in the control group. Between a pretest and posttest at the beginning and end of a 14-week semester, students in both the treatment and control groups completed a repeated listening exercise twice each week with novel recordings. Treatment group students participated in repeated listening exercise including a sequence of listening without the transcript, listening with the transcript, and finally listening again without the transcript. They answered comprehension questions between the first and second listening. Students in the control group listened to the same audio repeated the same number of times and answered the same comprehension questions but never had access to the transcript. Repeated measures ANOVA results for this quasi-experimental study show that the treatment group students outperformed the control group students on listening comprehension test items between the pretest and posttest. These findings suggest the potential benefit of transcript use to improve overall listening comprehension for advanced ESL learners. We encourage TESOL practitioners to consider whether transcript use might facilitate listening comprehension in their classrooms, and we call for additional research to determine the appropriateness of transcript use in other TESOL contexts.

Keywords

Listening comprehension, transcripts, reading while listening, post-listening activities

1 Introduction

Most scholars agree that listening is vitally important for second language learners (e.g., [Kurita, 2012](#); [Nunan, 2002](#); [Renukadevi, 2014](#)). Nevertheless, listening is often neglected compared to other

*Corresponding Author. Email: james_hartshorn@byu.edu

language skills (Vandergrift & Goh, 2012) and practitioners tend to feel least prepared to teach listening compared to other skills despite the fact that listening is often perceived as most important in academic settings (Hartshorn et al., 2019). We find it unacceptable that most practitioners feel the least prepared to effectively teach one of the most important language skills. As a field, we need to do much more to empower practitioners with effective training regarding the best ways to help their students develop their listening skills.

Rather than merely having students answer comprehension questions or discuss listening topics, Vandergrift (2007) has emphasized the value of a multi-stage listening lesson with a post-listening phase that allows students to discover what was actually said, asserting that this verification stage is essential to building future listening competencies. Many teachers may not consider how they might give such feedback on the mechanics of listening performance. They also may wonder how to manage feedback on learner ability to accurately segment the speech stream and recognize the aural forms of lexical items they know. One possible solution to providing students with the feedback suggested by Vandergrift (2007) could be through the use of written transcripts during a verification stage in repeated listening. Despite the vital importance of listening in language skill development, research on the effects of transcript use in repeated listening is quite scant. Much more research on the use of transcripts is vital if we are to empower practitioners with greater understanding and practical insights for the classroom.

Moreover, there are three prominent problems that have left considerable gaps in the literature and in our understanding about transcript use in the development of listening skills. The first is that much of the available research has been conducted over very short periods of time despite the fact that language development occurs gradually over many months and years. If a study is not of sufficient duration, it may be impossible to observe and measure aspects of language development that can only be seen over longer periods of time. The second prominent gap in the literature is the fact that many studies have neglected to account for proficiency differences among students. Finally, studies have been inconsistent in terms of when, how, and why transcripts have been used, making definitive claims about transcript use impossible. These kinds of limitations raise considerable concerns regarding the ecological validity of such studies and their findings.

Therefore, this study sought to address these three issues within an ecologically valid classroom context:

1. examining the effects of transcript use over the course of a 14-week semester—a longer duration than many similar studies,
2. properly account for proficiency level with a particular interest in the advanced proficiency level, and
3. describe thoroughly how transcripts were used along with the rationale for these discussions about transcript use.

The findings of this study should be of great interest to language practitioners, administrators, and researchers working toward improving the teaching and learning of vitally important listening skills in English.

2 Literature Review

2.1 The importance of listening instruction

At the outset, we need to clarify what we mean by the notion of listening. The aim of listening is the comprehension of spoken language. However, Goh (2014) has pointed out “listening is not just hearing,” and that the process may start “before the first speech signal is recognized, and it may go on long after the input or spoken information has stopped (p. 72). Listening is achieved through a series of highly complex and integrated “cognitive processes of perception, parsing, and utilization, which can be controlled and modified by learners through an enhancement of their knowledge of the language system, discourse, and how language is used” (Goh, 2014, p. 86).

The skills required for successful listening are critically important for students learning a second language (e.g., Kurita, 2012; Nunan, 2002; Renukadevi, 2014), particularly for learners who may be transitioning from an intensive English program to university study where English is the medium of instruction. For example, Hartshorn et al. (2017) surveyed professors from universities in the United States hosting the largest numbers of ESL learners and who taught in the most popular disciplines for ESL students. These scholars observed that to these content professors, listening and reading were the most important language skills for students beginning their academic careers. Other research has found similar results, suggesting that TESOL practitioners viewed listening and reading as the most important skills their students needed to master (Hartshorn et al., 2019).

Given the great importance of listening, it seems incongruous that many TESOL practitioners feel the least prepared to effectively teach the skill of listening compared to other skills such as reading, writing, speaking, and other areas of language development such as grammar, pronunciation, and vocabulary (Hartshorn et al., 2019). Vandergrift and Goh (2012, p. 4) have observed that despite elevated commitment to other language skills, “the development of listening receives the least systematic attention from teachers and instructional materials” and that learners are frequently “left to develop their listening abilities on their own with little direct support from their teachers.” Thus, it seems clear that research is needed to help us understand the best ways to teach listening, and that those insights need to be successfully transferred to TESOL practitioners.

2.2 How listening differs from reading

In the meantime, many TESOL practitioners may lack a sufficient understanding of the complexities involved in the skill of listening in a second language. Some assume that repurposing instructional methods, strategies, or activities that might be used in a reading class may be adequate. Like listening, reading comprehension is based on highly complex cognitive processes where readers construct a mental representation of a written text involving “word identification, parsing, referential mapping, and inference” which in turn interact “with the reader's conceptual knowledge” to create meaning (Perfetti & Adlof, 2012, p. 3). Though listening and reading share notable similarities as receptive skills that involve complex processing, practitioners need to understand the ways in which listening is distinct if they are to effectively help learners develop their listening skills.

In contrast to reading where the learner controls the speed, listeners cannot control the speech rate of interlocutors (Danan, 2016). Moreover, speech is usually produced as a “continuous stream” of sounds that need to be segmented and processed by the listener (Kim et al., 2012, p. 509) usually requiring “extensive practice” Huljstin (2003, p. 419) whereas in written form, word boundaries in a text are obvious. Flowerdew and Miller (2005) have noted that in reading texts, learners are likely to encounter more nominalization, subordination, complex coordination, passive voice, and attributive adjectives. They also tend to see much more precise vocabulary and a wide array of cohesive devices. On the other hand, Flowerdew and Miller (2005) have indicated that learners listening to spoken text will likely confront many more false starts, sentence fragments, unexpected pauses, as well as disconnected imperatives, questions, and references without sufficient context. Learners need to grapple with speech reductions and much more widely ranging expressions of register in spoken language than might be found in most written texts. Practitioners should thoughtfully consider the many ways in which listening is unique as they prepare classroom activities and work to help students develop their listening skills.

2.3 Factors influencing listening comprehension

The primary goal for teaching and learning associated with listening should be comprehension. Rost (2016) has observed the complex processes activated with listening. At the neurological level, processing

is influenced by the proper function of all the structural mechanisms tied to hearing and the appropriate allocation of attentional resources. At the linguistic level, processing is impacted by sound perception, the recognition of language units, the parsing of syntax, and the incorporation of nonverbal information. At the semantic level, processing is influenced by knowledge of schemata, linguistic structure, and the learner's ability to synthesize input, make appropriate inferences, and effectively address threats to comprehension. Finally, at the pragmatic level, processing includes the listener's ability to correctly infer intent, filtering the message through social expectations, regulating affective implications, and formulating appropriate responses. Since effective listening will necessarily utilize each of these processes, practitioners should consider their implications for specific teaching and learning contexts.

Other scholars such as Vandergrift and Goh (2012) also maintain that an important part of improving listening needs to include strategic use of metacognition. They explain "These metacognitive learning activities should aim to deepen learner understanding of themselves as L2 [second language] listeners, raise greater awareness of the demands and processes of L2 listening, and teach learners how to manage their comprehension and learning" (p. 13). Vandergrift and Goh have stressed affective undercurrents that may impact the development of listening skills. These could involve confronting anxiety related to listening, promoting self-efficacy by ensuring learners have successful listening experiences, and boosting motivation by assisting learners to appropriately apply strategies as they refine their listening skills.

Rost and Wilson (2013) have consolidated some of this thinking with five theoretical frameworks. These include the affective frame, emphasizing the motivation of the learner; the top-down frame, targeting improved understanding and interpretation; the bottom-up frame, focusing on the perception of sounds, the recognition of words, and the parsing of language; the interactive frame, highlighting the collaboration and interdependence needed in listening; and the autonomous frame, accentuating the strategies needed for effective development of listening skills. The hope of Rost and Wilson is that by learning about and keeping in mind these frameworks, practitioners may be better able to develop a variety of listening activities appropriate to the needs of their own students.

Practitioners who understand and effectively apply the preceding information about the many practical and theoretical considerations involved in ESL listening pedagogy are the most likely to help their students to successfully develop their listening skills. Yet, given the great importance of the skill of listening in a second language and the many complex factors that tend to make it challenging, ongoing research is vital to help improve our understanding of the best ways to facilitate teaching and learning.

2.4 Using transcripts for listening development

In addition to evidence that suggests that repeated listening can improve comprehension (Rodrigo, 2017; Šendaž et al., 2018), some scholars have suggested there may be benefits to using written transcripts or "the written versions of listening material" (Brown & Brown, 2011, p. 13) during listening instruction and practice. While some have focused on the use of captions for video use (e.g., Gruba, 2004; Winke et al., 2010), others have examined more traditional listening practice with written transcripts (e.g., Danan, 2016; Vandergrift, 2011; Vandergrift & Goh, 2012). Brown and Brown (2011) have described transcript use as "controversial" (p. 13). This is because traditional thinking has suggested that such an aid could create dependency for the learners and foster casual and ineffective approaches to listening. Though Brown and Brown have acknowledged these as possible detriments, they also argue that the classroom needs to facilitate language development which may be aided by the Vygotskian concept of utilizing appropriate scaffolding (Kozulin, 1998). Lund (1991) has suggested that while transcripts probably should not be used every day, strategic use of transcripts might help students make better sense of the language they might otherwise miss and could prepare them for more successful listening in the future.

Some scholars have found a number of benefits from using transcripts. For example, Chang (2011) observed that using transcripts of spoken texts, or reading-while-listening (RWL), helped her students

improve in their segmentation of word boundaries. It also helped them acclimate to speech rate, language chunking, and the rhythmic flow of the speech texts they encountered. Danan (2016) noticed that textual support helped novice learners to keep up with the listening materials and to recognize more of the words they encountered. These benefits seemed to persist even after the transcripts were later removed. Advanced learners benefited most from transcripts in terms of improved comprehension of specific details. On the other hand, the benefits of transcript use for intermediate learners was much more idiosyncratic depending on specific language learning needs and learning style preferences.

Other studies on reading that use the RWL protocols may also be highly relevant. For example, Teng (2018) who used a read-only and a RWL group found that the RWL group incidentally learned more vocabulary than the read-only group. Similarly, Conklin et al. (2020) used eye tracking to examine L1 and L2 learners engaged in reading-only tasks and RWL tasks. In the reading-only tasks, the L2 learners read more slowly with longer fixations compared to the L1 readers. When using the audio and text together, L2 learners were closer to that of their native-speaking counterparts. The researchers noted that using the audio and text together helped L2 students to “segment, decode, parse and/or make the form-meaning link for words in the text” (p. 273).

Though L2 learners lagged behind the text more than the L1 students, the eye movements of both groups generally preceded the text. This may seem counterintuitive since the visual stimuli of the text was being processed just before the auditory signal. Nevertheless, researchers suggested that by looking just ahead of the text while listening may allow the learners to more effectively link the phonological form and the orthographic form with the meaning. Thus, the use of transcripts while completing listening tasks might not only aid comprehension for the immediate task but may facilitate important aspects of language development.

2.5 Questions about transcript use

Despite the potential benefits of transcript use shown across some studies, not all research has demonstrated a clear benefit from transcript use. Diao et al. (2007) found that students with a transcript received substantially higher scores for listening comprehension questions, but when students were tested again a week later without a transcript, there was no apparent benefit. The authors concluded that while the transcript helped with comprehension, it did not help students “in the construction of relevant schemas” (p. 250) needed for successful listening comprehension. Despite this conclusion, since the development of listening skills takes time, using transcripts on just one occasion may not have been adequate to show the effects of transcript use over a longer period.

Focusing on cognitive load theory, Roussel et al. (2019) sought to identify whether listening comprehension improved more from top-down processing (e.g., use of context cues, application of schemata, drawing inference, predicting) or bottom-up processing (e.g., sound recognition, segmenting and parsing words and syntax). Students in the two groups performed similarly on a pretest and then the two groups received three weeks of top-down or bottom-up training. Though overall there were no significant differences in listening comprehension across groups, they did find meaningful gains for those in the bottom-up group who produced the lowest scores on the pretest. Subsequently, the researchers concluded that reducing the cognitive load for lower-level listeners through bottom-up training may be useful.

Some scholars have invoked cognitive load theory to the notion of RWL. For example, Moussa-Inaty, et al. (2012) conducted research over a three-day period that sought to determine whether the addition of reading a transcript while listening adversely adds to the cognitive load of the listener or whether the additional reading of the transcript may reduce the cognitive load on the listener. They examined three groups including listening-only, reading-only and RWL. They observed that the students they studied demonstrated lower levels of listening comprehension during RWL and that the best performance levels

were from the reading-only group. However, language proficiency was not clearly assessed, the term of the study was excessively short, and transcripts were used as a during-listening activity rather than used as scaffolding as a post-listening activity.

Despite these conclusions, we are concerned that Roussel et al. (2019) may have conflated cognitive load with the mode of training (top-down or bottom-up). More importantly, we worry that three days (Moussa-Inaty, et al., 2012) or even three weeks (Roussel et al., 2019) may still be too short to observe meaningful benefits for language development. Language learning is complex and often requires substantial time and energy. Skill acquisition theory suggests that with adequate instruction and abundant practice and feedback over an extended period of time, language use will become more automatic reducing the cognitive load on the learner (Dekeyser, 2020).

Effectively addressing many of the research design limitations of previous studies, Sohler (2020) sought to further examine the potential benefits of RWL. She taught two intermediate-level listening and speaking classes of 17 students each who completed a 14-week course. One was designated as the RWL treatment group and the other was designated as the control group. Both groups engaged in work with two passages each week, taking about thirty to forty minutes of class time each week. Both groups listened to the passages three times. The control group listened without transcripts while the treatment group was provided with a transcript on the second listening (but not for the first or third listening). Both groups made substantial gains in listening comprehension over the course of the semester. The treatment group slightly outperformed the control group. Though technically the difference was not statistically significant ($p=.101$), the difference produced a medium effect size ($\eta^2_p = .082$) suggesting a meaningful benefit to the treatment group compared to the control group.

2.6 Drawing from the literature and classroom observation

In addition to the preceding literature about the complexities and intricacies associated with the teaching and learning of listening, we consistently noticed the following observations from our own classroom contexts.

1. Many students have continued to struggle with comprehension despite extensive work with listening strategies and a great deal of practice.
2. As students have written what they thought they heard while engaged in dictation exercises, many have revealed that they misheard a great deal of key information.
3. Many students have not been able to recognize words due to pronunciation features such as reductions, dialectal variations, sentence stress patterns, and challenges segmenting the speech stream.
4. Students have often experienced moments of enlightenment when they see transcripts of texts and suddenly understand what they have listened to but had not been able to fully decipher without the transcript.

Drawing from these personal observations from our own classroom experiences and the preceding review of literature, we note several important insights. First, we see that the language skill of listening is extremely complex but vitally important. Many consider it one of the most important language skills learners need, particularly in academic settings. Yet, practitioners tend to feel the least prepared to teach listening effectively. This suggests that the field must do much more to empower practitioners with the methods and insights needed to help them successfully facilitate the development of listening skills. One of the most important ways to do this is through targeted research that provides the very insights that are needed.

Second, we note that research on listening with transcripts is scant and incomplete. This review of the literature has shown possible benefits for using transcripts to assist listening development. It has also shown mixed results across studies that have utilized transcripts. Broad claims about language

acquisition phenomena should not be based on just a few studies or even our own classroom experiences. Much more research is needed to create a critical mass of findings that will help us better understand language development and how to leverage this insight in the classroom. This is particularly true if the limitations of previous studies have inaccurately skewed our perceptions. The literature has revealed three major challenges with a number of previous studies. First, this includes studies whose durations may not have been long enough to accurately capture the longer-term effects of transcript use over time. Second, this includes studies where either language proficiency was not accounted for, or where there are broad claims that go beyond the specific proficiency levels included in the study. Additionally, there has been inconsistency across studies in terms of when, how, and why transcripts should be used in the listening class.

Given these findings, this study seeks to address these three specific gaps in the literature and to add to our understanding of the effect of transcript use on the development of listening. This study will specifically examine the effect of transcript use in the verification stage (Vandergrift, 2007) of a listening lesson over the course of one full 14-week semester. Based on the preceding literature and classroom observation that suggested the possible benefits of using transcripts to enhance listening practice, we determined to examine whether transcript use could help facilitate listening development. Since intermediate student gains in listening comprehension were not significantly better than the control group in the Sohler (2020) study, this study was specifically designed to examine advanced English language learners rather than intermediate proficiency students to determine whether language proficiency may also be a factor in the effect of transcript use.

3 Research Question

Given the preceding discussion, the following research question was articulated: Does repeated listening with transcript use at the verification stage of a listening lesson improve comprehension for advanced proficiency students more than repeated listening without a transcript?

4 Methods

This section describes the methods used to conduct this research including a description of the participants, the treatment, data elicitation, and the planned analyses.

4.1 Participants

4.1.1 Students

This study examined the listening comprehension of 31 students who were studying English as a second language in an intensive English program (IEP) in the western United States. Students were tested and placed into the highest level of the IEP (advanced-low to advanced-mid). This was done by interviewers who have been well trained and certified to conduct and rate oral proficiency interviews according to the guidelines established by American Council of Teachers of Foreign Languages (ACFTL, 2012), and who had participated in extensive calibration practice (McNamara, 1996). Moreover, fair averages were used from Many Facets Rasch Modeling to adjust for any slight rater biases (e.g., Eckes, 2011).

Of the 31 students included in the study, 16 were randomly assigned by the institution to one class and 15 were randomly assigned to the other class. Since this was done prior to the study and outside of the purview of the researchers, we are treating these groups as intact classes. We randomly selected

one class as the treatment group (15 students) and the other as the control group (16 students). Table 1 summarizes the native language and gender composition of both groups. The major native language group in both classes was Spanish. With respect to males and females, the control group was equally divided though the treatment group was slightly less balanced.

Table 1

Experimental Groups by Native Language and Gender

Native Language	Experimental Groups					
	Control Group			Treatment Group		
	Male	Female	Total	Male	Female	Total
Spanish	5	5	10	4	7	11
Chinese	2	1	3	0	2	2
Portuguese	1	0	1	0	1	1
Korean	0	0	0	0	1	1
Mongolian	0	1	1	0	0	0
Swedish	0	1	1	0	0	0
Totals	8	8	16	4	11	15

4.1.2 The Teacher

In an attempt to minimize teacher effect, the same teacher taught the students in both the treatment and control group classes where listening was a primary focus of the course (Jackson et al., 2014). Though students in both groups experienced the same curriculum, they did not have the same teachers for their other three courses in the program including Reading, Writing, and Grammar. The teacher who taught the treatment and control classes held a TESOL MA degree and had about five years of experience teaching in an IEP.

4.2 Procedures

Great effort was expended to minimize potentially confounding variables to help ensure, to the extent possible, that any observed effects were the result of the treatment itself. These efforts included the following for both the treatment and control groups:

1. Students were randomly assigned to the treatment and control groups.
2. Pretests and posttests were identical for both groups and data were collected at the same time at the beginning and end of the semester for both groups.
3. Students in both groups were taught at the same institution, during the same semester, and by the same teacher (with the hope of eliminating teacher effect).
4. All classes in which the treatment and control protocols occurred were held in the early afternoon on the same days, Monday through Thursday.
5. With the exception of the treatment, students were taught with the same curriculum. This included the same Listening and Speaking class curriculum as well as the same curriculum for their other three classes taken by all students (Reading, Writing, and Grammar)

With the intent of strengthening the reliability of the elicited data, the pretest and posttest results were derived from data from more than one listening passage on different topics (e.g., Shadish, Cook, & Campbell, 2002). In the first week of class, the pretest included student responses to ten objective items

for two separate listening exercises. Scores were averaged for the two listening passages. For each pretest component, students listened to an audio file one time without any introduction or warmup activity. After each audio file was played, students had five minutes to answer the accompanying questions. The two listening exercises and accompanying questions on the pretest took approximately 15 minutes to complete. Because the researchers intended to repeat the audio files from the pretest as part of the posttest data collection, the teacher did not review answers or use that audio content again during the semester.

Though we recognized the potential benefits of a delayed posttest after the completion of the treatment, our specific intent was to extend the treatment as long as possible during the semester, particularly since many previous studies were of such short duration. Knowing we may not have additional access to the students after the end of the semester, we therefore decided not to include a delayed posttest to ensure that students experienced a full semester of their listening instruction and practice.

Subsequently, the posttest was given during week 14 and included student responses to twenty objective items for four separate listening exercises. For the posttest, the original recordings and questions were used in addition to two more novel audio files that were at the same level of difficulty. All of the audio materials were comparable in length, difficulty, and topic. Transcripts were not shown to students at any time during the pretest or posttest. The four listening exercises and accompanying questions included on the posttest took about 30 minutes to complete.

4.2.1 Treatment group

Students randomly assigned to the treatment group participated in a podcast listening activity twice each week for approximately 15 minutes on each occasion. These activities continued from week 2 through week 13. The podcast series *60-Second Science from Scientific American* was chosen for several reasons. First, while the podcasts are not all limited to 60-seconds, they are short enough to lend themselves well to repetition without taking up a great deal of class time. In addition, the podcast episodes are similar in difficulty and were perceived to be at the ideal proficiency level—just beyond what these advanced learners would easily understand on their own. It was assumed that this would allow some room for improvement between the pretest and posttest. Moreover, the content of the podcast material seemed to be current, authentic, and highly engaging. Finally, the transcripts for each of the podcasts were available on the website facilitating transcript use.

In the classroom, the teacher simply introduced the podcast to the treatment group by stating the topic. No vocabulary was pre-taught and no background knowledge was intentionally activated by the teacher. The 15-minute procedure included four steps.

1. *First listen.* For the students' first exposure to a particular podcast, they just listened without a specific task to complete such as taking notes.
2. *Comprehension Questions.* The students were led through an activity of informally answering and discussing four multiple-choice comprehension questions written by the teacher to address learning objectives from the class (e.g., identify the main idea, identify major details, etc.). The teacher displayed the questions one at a time on a PowerPoint slide. After giving the students a minute to decide on their answers, students reported their answers either verbally or by raising their hands. After each question was presented to the students, the teacher stated the correct answer and reinforced strategic knowledge that was needed to answer each question.
3. *Second listen (with transcript).* After the comprehension questions were finished, the teacher distributed a copy of the transcript for the podcast they had just listened to. Students were instructed to listen again and follow along with their eyes. Special emphasis was given to matching the aural and visual form of the words simultaneously, rather than reading through the transcript more quickly or more slowly than the pace of the speaker and ignoring the repetition of the audio. Either as they

listened or right after, students marked any words they didn't recognize the first time the audio was played. Students were told not to discriminate between words they knew (but didn't recognize) and words that were unfamiliar. All of the unrecognized words were marked. Students then had a few minutes to ask questions about lexical items they had not understood or look up new vocabulary words they found in the transcript.

4. *Third listen (without transcript)*. Students listened to the podcast for a third time, specifically listening for what they had not recognized during their first encounter with the podcast.

4.2.2 Control group

Students randomly assigned to the control group listened to the same podcasts on the same days as those in the treatment group during weeks 2 through 13. To isolate the effect of the transcript with the treatment group, the control group answered the same comprehension questions and also listened to the podcast three times. The teacher introduced the podcast in the same way to both the treatment and control groups. Then students in the control group listened to the podcast twice, answered the comprehension questions in the same manner as the treatment group, and then listened again to identify answers to the comprehension questions they were not able to determine the first two times they listened. Table 2 summarizes the procedures used for both groups.

In an effort to identify possible differences for the effect of the listening activities and teacher impact on the respective groups, students were asked at the end of the semester to complete a seven-point survey item (ranging from strongly agree to strongly disagree) about their perceived benefit from the activities. Interestingly, mean responses for both groups were exactly the same with no difference, $t(6)=0.00$, $p=1.00$ ($M=6.75$, $SD=.50$), indicating a mean response approaching *strongly agree*. This suggests that both groups of students believed they benefited from the course and the teacher's efforts to help them develop their listening skills.

Table 2

Summary of Procedure for both Treatment and Control Groups

Step	Treatment Group	Control Group
Step 1	Listen to the podcast once (without transcript).	Listen to the podcast twice (without transcript).
Step 2	Answer and discuss comprehension questions.	Answer and discuss comprehension questions.
Step 3	Listen (with transcript). Mark unrecognized words and phrases. Allow time to investigate lexical items.	(not included)
Step 4	Listen to the podcast third time (without transcript).	Listen to the podcast third time (without transcript).

4.2.3 Instruments

For both the pretest and the posttest, students listened to two types of recordings and answered accompanying multiple-choice comprehension questions. The first type was the podcast used in class throughout the semester from *Scientific American*. The other type included TOEFL iBT test content. While TOEFL test content was not used in classes during the semester, it was chosen as part of the

pretest and posttest because it utilized appropriate questions and had been tested for reliability. The expectation was that using the TOEFL iBT content would strengthen the reliability and validity of the pretest and posttest.

4.3 Data analysis

This quasi-experimental study utilized a mixed-model repeated measures analysis of variance. This was chosen to account for a between-group factor (the treatment versus the control group) as well as a within-subjects factor (repeated measure from pretest to posttest) (see Maxwell & Delaney, 2004). The posttest was made up of the averages of student responses from four separate listening exercises designed to be similar to the pretest. In an effort to make the composite posttest relatively equivalent to the pretest, it included newly collected data from the two instruments used in the pretest (on the topics of hurricanes and frogs) as well as two additional instruments that were designed to be similar to the pretest instruments—one about climate and one about hummingbirds. These were all taken in the last week of class. A summary of these test components is illustrated in Table 3.

Table 3

Composite Pretest and Posttest Components

Elicitation	Treatment	Control
Pretest	Hurricanes (Podcast)	Hurricanes (Podcast)
	Frogs (TOEFL)	Frogs (TOEFL)
Posttest	Hurricanes (Podcast)	Hurricanes (Podcast)
	Frogs (TOEFL)	Frogs (TOEFL)
	Climate (Podcast)	Climate (Podcast)
	Hummingbirds (TOEFL)	Hummingbirds (TOEFL)

5 Results

Before examining the analysis designed to answer the actual research question, we review the data itself to ensure it is suitable for analysis. Raw scores from the pretest and posttest were converted into ratios ranging from 0 to 1, reflecting the overall accuracy of student listening comprehension. Pretest performance levels for the students in the control ($M=.425$, $SD=.198$) and treatment ($M=.436$, $SD=.172$) groups were similar. No statistically significant difference was observed, $t(29) = .173$, $p=.864$, suggesting the comparability of student listening ability prior to the application of the treatment.

Repeated measures analyses of variance were also used with the posttest results for the control and treatment groups to determine whether performance level across the four posttest components were similar. For the control group, no statistically significant (Following guidelines from Field (2005), Greenhouse-Geisser corrections were used for the analyses of variance for the control and treatment groups due to a lack of sphericity resulting in adjusted degrees of freedom for the F-test results.) difference was observed across the four sets of scores that were combined to form the single posttest score, $F(1.590,20.673) = 1.745$, $p = .202$. Similarly, no statistically significant difference was observed for the four data sources for the treatment group posttest, $F(2.701,29.711) = 1.455$, $p = .248$. These findings suggest the relative comparability of the four posttest components and their appropriateness for this study.

For the control group, no statistically significant difference was observed between the instruments only used for the posttest including the listening passage about hurricanes ($M=.304$, $SD=.175$) and the listening passage about climate ($M=.429$, $SD=.385$), $t(13)=1.391$, $p=.187$. Similarly, for the treatment group, no statistically significant difference was observed between the posttest use of the listening about hurricanes ($M=.604$, $SD=.167$) and the listening about climate ($M=.710$, $SD=.334$), $t(11)=1.00$, $p=.339$. These results suggest the comparability of these instrument components and that the instruments functioned as intended.

Since this study was conducted in a formal educational setting specifically designed to help students to develop their language skills, it was assumed that students would make observable progress in the development of their listening skills over the course of the study. Nevertheless, we recognized that the most important result would be the group by time interaction effect showing whether or not the treatment group would outperform the control group from the beginning to the end of the study.

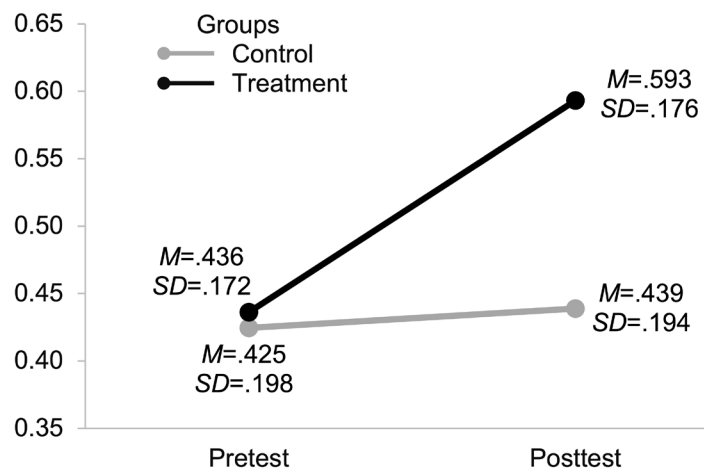
To answer the research question itself, a mixed model repeated measures ANOVA was used to examine the effect of the treatment on listening comprehension as measured by the composite pretest and posttest results described previously. This analysis produced three results. This includes the overall effect of time on student listening skill development from pretest to posttest without regard for group. It also includes the overall effect of group without regard to differences from the pretest to the posttest. However, the most important result for this study is the interaction effect of group over time, which should provide important information to answer our research question. If the interaction effect of group by time is significant, it will contextualize and supersede the other main effects of time or group.

As should be expected in a formal setting of teaching and learning, the collective group of all students combined demonstrated statistically significant progress in their listening $F(1,29) = 9.554$, $p = .004$, $\eta^2_p = .248$. This suggests that student progress for listening was observed without regard for teaching method. On the other hand, group differences without regard to time were not statistically significant, $F(1,29)$, 1.865 , $p = .183$, though a medium effect size was observed favoring the treatment group, $\eta^2_p = .060$.

Despite these results, however, the key analysis here is the group by time interaction. Although no statistically significant difference was observed between groups at the pretest as shown previously, the group by time interaction was statistically significant, $F(1,29)= 6.630$ $p=.015$, and produced a large effect size, $\eta^2_p = .186$. This interaction is illustrated in Figure 1 which presents listening ratio scores on the vertical axis and shows group means (M) and standard deviations (SD) for the pretest and posttest scores. These results show that while the improvements in listening comprehension for the control group were nearly negligible, they were quite substantial for the treatment group.

Figure 1

Treatment Effect on Listening Comprehension



The goal of this research was to determine the effect of repeated listening activities with transcripts used during post-listening to provide students with an opportunity to verify what they heard versus what they did not. These findings suggest that regularly incorporating transcript use in the verification stage of a listening lesson may help advanced proficiency students improve their listening comprehension, even when transcripts are not used in novel listening tasks. These results seem consistent with findings from scholars such as Teng (2018) and Chang (2011) where students who engaged in RWL outperformed students in control groups.

Compared to previous studies, this finding may be attributed to a variety of factors. Where previous research used various approaches to transcript use during the first time an audio was heard, in this study, transcripts were only used during the post-listening verification phase of the lesson. Though using transcripts from the outset of listening practice could foster detrimental dependency on transcripts (e.g., Brown & Brown, 2011), having students exert themselves first while listening without the transcript may have useful benefits. First, it could help them develop greater awareness that there are gaps in their comprehension. Second, that realization may help prepare students with greater learner readiness and motivation to notice and learn from the transcript when they eventually see it in the verification stage of the lesson (e.g., Aguilar & Kim, 2019; Bozkurt & Arslan, 2018; Kartal & Balcikanh, 2019).

In addition, learners observed in this study completed the repeated listening activity twice a week for an entire semester. Though not a substantial amount of the course, this consistent pattern was more frequent and of longer duration than interventions used in many other studies. Although Sohler (2020) used a similar repeated listening activity, the intermediate proficiency of her learners, her tool, and her audio texts varied in significant ways from this study. Much more research may be needed to determine specifically which proficiency levels may benefit the most from transcript use in listening classes.

Given the significant findings of this study, the eye tracking research by Conklin et al. (2020) cited previously becomes even more compelling, particularly in how transcript use seemed to help L2 learners perform more like their L1 counterparts. The fact that transcript use also seemed to help L2 learners link phonological and orthographic forms with meaning is of great interest in terms of overall language development.

Some of the research cited previously attempted to address the question of cognitive load when combining transcript use with listening activities. Though an examination of cognitive load was not a direct focus of our research in this study, given our positive results with advanced learners, we now believe it may be important to consider cognitive load in future research. Nevertheless, we question whether an examination of the effect of transcripts on cognitive load may oversimplify the phenomenon if language proficiency is not also taken into account. Cognitive load theory described by Pillay (1994) considers the mental effort needed to successfully complete a task with the expectation that learning is likely to be more successful with the cognitive load is minimized. Nevertheless, skill acquisition theory suggests that with the increased language automatization that comes with improved language proficiency, cognitive load will be reduced (Dekeyser, 2020).

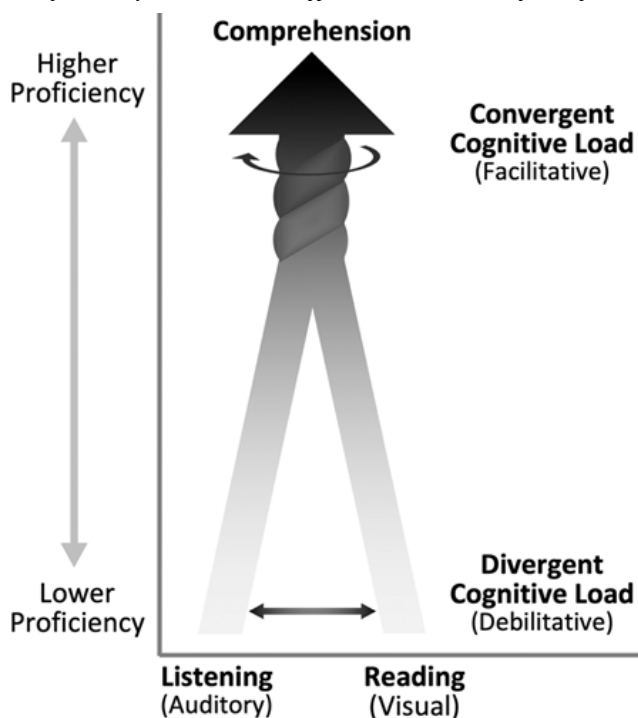
Since language proficiency is likely to have a substantial impact on cognitive load, we suggest that proficiency level may be an important consideration in future research of the effects of transcript use on listening comprehension. This suggestion seems consistent with the findings of this study which produced a large effect size from transcript use with advanced learner, compared with the findings of the Sohler (2020) study which were not statistically significant though they produced a medium effect with intermediate students.

In considering the results of this study along with the other cited literature, we wonder whether transcript use may have a facilitative or debilitating effect on listening depending on the language proficiency level of the student. In an effort to guide future research, we hypothesize that transcript use at lower proficiency levels may either interfere or have no appreciable impact on learning while effective use of transcripts at higher proficiency levels may facilitate the development of listening comprehension.

To help clarify this hypothesis, Figure 2 depicts listening and reading as two strands of a rope. At lower proficiency levels at the bottom of the figure, learners may be more likely to struggle to simultaneously process the separate phonological and orthographic information. Attempting to focus on one strand may exhaust the learner's attentional capacity to focus on the other strand. The divergent nature of attempting to process the separate strands may lead to cognitive overload and potentially undermine comprehension and language development.

Figure 2

Proficiency's Possible Effect on the Benefits of Transcripts in Repeated Listening



On the other hand, at higher proficiency levels toward the top of the figure, the convergence of the phonological and orthographic information may have a facilitative effect on language development where comprehension of the listening may be aided by the text and comprehension of the text may be aided by the listening. Of course, this hypothesis needs to be thoroughly examined through additional study. We therefore call for more research on transcript use with repeated listening that carefully accounts for language proficiency in ways that may help support or refute the role of proficiency level in transcript-based listening comprehension research. Future research could also examine whether lower proficiency learners could benefit from certain kinds of transcript use by managing the cognitive load.

In the meantime, appropriate use of transcripts in repeated listening activities may be a simple but useful tool for teachers who feel underprepared to teach listening skills. By using transcripts as post-listening feedback for students in the verification phase of a listening lesson, teachers may help learners to notice important features of the language and their listening experience that might be impossible for teachers to point out from observation alone.

6.1 Limitations and future research

This research includes several limitations that should be acknowledged to help readers to appropriately contextualize these findings and to help researchers to improve future study. First, this study only

included 31 students who were already divided into intact classes when the study began. Though pretest data showed comparable performance levels from both groups, a larger sample of students would be much more ideal. Further, the listening materials used in this study were from a single podcast series designed primarily for the United States. Though appropriate for the specific context of this study, it is uncertain whether or to what extent other kinds of listening materials might generate similar results. Moreover, about two-thirds of the students in this study were native speakers of Spanish. Each of these limitations could impact the generalizability of these findings. We recommend, therefore, that future research examine a greater variety of listening materials, a much larger sample of students across more L1s, and that it utilizes randomly assigned control and treatment groups.

It was the explicit intention of this study to examine the effects of transcript use for advanced proficiency students over the course of one 14-week semester. Thus, these findings should not be generalized for learners at other proficiency levels. Nevertheless, we recognize that including other proficiency levels in a larger study could have yielded additional insights about the effects of proficiency along with the effects of transcript use. Though this study was much longer than many of the previous studies cited in this paper, it did not include a delayed posttest. We recognize the need for more long-term studies, particularly those that include a delayed posttest to help identify the extent to which any observed benefits of transcript use might endure after the treatment has ended.

We also acknowledge and encourage other areas ripe for additional study. In particular, these include the specific examination of the effect of cognitive load with transcript use, an important consideration not explicitly addressed in the original design of this study but that we believe is worthy of future examination. We specifically call for further research designed to gather evidence that supports or refutes our hypothesis illustrated in Figure 2, that transcript use with the kind of listening activities described in this study may create a convergent cognitive load (facilitative) at higher proficiency levels while creating a divergent cognitive load (debilitative) in lower proficiency levels. We also encourage additional research in areas such as examining the effects of affect factors on listening, the effects of collaborative listening (Saeedakhtar, et al., 2021), and issues related to providing additional support for listening students, particularly those at lower proficiency levels (Zhang & Graham, 2020).

7 Conclusion

This study examined the effect of transcript use to support listening development for advanced ESL learners. Findings suggest that regular use of transcripts with repeated listening may facilitate listening skill development, even when transcripts are not used subsequently. This study specifically makes four important contributions to the literature. First, this study used an ecologically valid classroom context to examining the effects of transcript use over the course of a 14-week semester—a much long duration than many similar studies. Second, it properly accounted for language proficiency with a particular interest in the effect of transcript use of advanced proficiency students. Third, it described in detail how transcripts were used along with the rationale for these discussions. Finally, it presents a hypothesis regarding cognitive that sets an important research agenda for the future.

The findings of this study have important implications for TESOL practitioners and curriculum materials designers. When teachers better understand how to effectively utilize the post-listening phases of a listening lesson to provide students with feedback on actual listening performance (operationalized in this study through transcript verification), listeners may become more proficient at dissecting the speech stream and decoding the aural messages they encounter. Though more studies are needed to validate these results, these findings may help teachers, curriculum developers, and materials writers to better understand important ways to optimize the three phases of a listening lesson and provide more explicit opportunities to foster scaffolded skill development.

References

- ACTFL. (2012). *ACTFL proficiency guidelines 2012*. Alexandria, VA: American Council on the Teaching of Foreign Languages. https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf
- Aguilar, S. J., & Kin, A. Y. (2019). Hispanic students' sense of control in relation to postsecondary enrollment outcomes. *Social Sciences*, 8, 67.
- Bozkurt, N., & Arslan, F. Y. (2018). Learner autonomy in language learning: Syrian refugee EFL learners' perceptions and readiness. *Multidisciplinary Journal of Educational Research*, 8(2), 115-145.
- Brown, S., & Brown, S. R. (2011). *Listening myths: Applying second language research to classroom teaching*. University of Michigan Press.
- Chang, A. C. S. (2011). The Effect of Reading While Listening to Audiobooks: Listening Fluency and Vocabulary Gain. *Asian Journal of English Language Teaching*, 21.
- Conklin, K., Alotaibi, S., Pellicer-Sánchez, A., & Vilkaitė-Lozdienė, L. (2020). What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second Language Research*, 36(3), 257-276.
- Danan, M. (2016). Enhancing listening with captions and transcripts. *Applied Language Learning*, 26, 1-24.
- DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (3rd., pp. 83-104). Routledge.
- Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *The American journal of psychology*, 120(2), 237-261.
- Eckes, T. (2011). Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments. Peter Lang.
- Field, A.P., (2005). *Discovering Statistics Using SPSS*. London: Sage Publications.
- Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge university press.
- Goh, C. C. M. (2014). Second language listening comprehension: Process and pedagogy. In D.M. Brinton, M. Celce-Murcia, & M. A Snow (Eds.), *Teaching English as a second or foreign language* (pp. 72-89). Heinle Cengage Learning.
- Gruba, P. (2004). Understanding digitized second language videotext. *Computer Assisted Language Learning*, 17(1), 51-82.
- Hartshorn, K. J., Evans, N. W., Egbert, J., & Johnson, A. (2017). Discipline-specific reading expectation and challenges for ESL learners in US universities. *Reading in a Foreign Language*, 29, 36-60.
- Hartshorn, K. J., Hart, J., & McMurry, B. (2019). Comparing language skill priorities among TESOL faculty and ESL students bound for English-medium universities. *TESOL Journal*, 10(3), 1-18.
- Hulstijn, J. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *Computer Assisted Language Learning*, 16(5), 413-425. https://pure.uva.nl/ws/files/1930350/165026_Hulstijn_CALL_2003.pdf
- Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annu. Rev. Econ.*, 6(1), 801-825.
- Kartal, G., & Balcikanlı, C. (2019). Tracking the culture of learning and readiness for learner autonomy in a Turkish context. *TEFLIN Journal*, 30(1), 22-46.
- Kim, D., Stephens, J., & Pitt, M. (2012). How does context play a part in splitting words apart?

- Production and perception of word boundaries in casual speech. *Journal of Memory and Language*, 66, 509-529. <https://doi.org/10.1016/j.jml.2011.12.007>
- Kozulin, A. (1998). *Psychological tools: A sociocultural approach to education*. Harvard University Press.
- Kurita, T. (2012). Issues in second language listening comprehension and the pedagogical implications. *Accents Asia*, 5(1), 30-44.
- Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *The Modern Language Journal*, 75(2), 196-204.
- Maxwell, S. E. & Delaney, H. D. (2004). *Designing experiments and analyzing data* (2nd ed.). Lawrence Erlbaum Associates.
- McNamara, T. F. (1996). *Measuring second language performance*. New York: Longman.
- Moussa, I. J., Ayres, P., & Sweller, J. (2012). Improving listening skills in English as a foreign language by reading rather than listening: A cognitive load perspective. *Applied Cognitive Psychology*, 26(3), 391-402. <https://doi.org/10.1002/acp.1840>
- Nunan, D. (2002). Listening in language learning. *Methodology in language teaching: An anthology of current practice*, 238-241.
- Perfetti, C. A., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. In J. P. Sabitini & E. R. Albro (Eds.), *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences* (pp. 3-20). Rowman & Littlefield Education.
- Pillay, H. K. (1994). Online reading: a preliminary study of the impact of integrated and split-attention formats on L2 students' cognitive load. *Instructional Science*, 22(2), 91-113.
- Renukadevi, D. (2014). The role of listening in language acquisition; the challenges & strategies in teaching listening. *International journal of education and information studies*, 4(1), 59-63.
- Rodrigo, V. (2017). Quantifying comprehension gains after repeated listening by students of Spanish with different listening ability: an exploratory study. *Journal of Spanish Language Teaching*, 4(1), 35-48.
- Rost, M. (2016). *Teaching and researching: Listening* (3rd ed.). Routledge.
- Rost, M., & Wilson, J. J. (2013). *Active listening*. Routledge.
- Roussel, S., Gruson, B., & Galan, J. P. (2019). What types of training improve learners' performances in second language listening comprehension?. *International Journal of listening*, 33(1), 39-52.
- Saeedakhtar, A., Haqju, R., & Rouhi, A. (2021). The impact of collaborative listening to podcasts on high school learners' listening comprehension and vocabulary learning. *System*, 101, 102588.
- Şendağ, S., Gedik, N., & Toker, S. (2018). Impact of repetitive listening, listening-aid and podcast length on EFL podcast listening. *Computers & Education*, 125, 273-283.
- Shadish, W. R., Cook, T. D., Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Sohler, S. (2020). *Developing Listening Comprehension in ESL Students at the Intermediate Level by Reading Transcripts While Listening: A Cognitive Load Perspective*. Brigham Young University.
- Teng, F. (2018). Incidental vocabulary acquisition from reading-only and reading-while-listening; a multidimensional approach. *Innovation in Language Learning and Teaching*, 12(3), 274-288.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language teaching*, 40(3), 191-210.
- Vandergrift, L. (2011). Second language listening. *Handbook of research in second language teaching and learning*, 2, 455.

- Vandergrift, L., & Goh, C. C. (2012). Teaching and learning second language listening: Metacognition in action. Routledge.
- Winke, P., Gass, S., & Syodorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65-86.
- Zhang, P., & Graham, S. (2020). Learning vocabulary through listening: The role of vocabulary knowledge and listening proficiency. *Language Learning*, 70(4), 1017-1053.

K. James Hartshorn received his PhD in instructional psychology with a specialization in second language acquisition and an ESL emphasis. His involvement in second language education in United States and Asia spans more than three decades. He currently serves as a faculty member in Brigham Young University's Linguistics Department. His professional interests include mentoring new professionals in the TESOL field by helping them to optimize the efficacy of their teaching and research.

Chirstin Stephens currently serves as the Coordinator of Curriculum and Assessment in the American Language and Culture Center at Southern Utah University. She specializes in curriculum and materials development, placement, assessment, teacher training and support, and program evaluation. She is particularly interested in helping English language learners successfully transition to university study in English-medium universities.