The Effects of Real-life Listening Practice on ESL Students'

Listening Comprehension Improvement

Vadym Malyshkevych

A project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Dr. Lynn E. Henrichsen, Chair
Dr. Dee I. Gardner
Dr. K. James Hartshorn

Department of Linguistics and English Language

Brigham Young University

ABSTRACT

The Effects of Real-life Listening Practice on ESL Students'
Listening Comprehension Improvement
Vadym Malyshkevych

Department of Linguistics and English Language, BYU
Master of Arts

This project is an attempt to detect if a cause-effect relationship exists between real-life listening practice by Advanced Low ACTFL level ESL students and their progress in listening comprehension. The underlying idea of this study is based on J.R. Anderson's classic model of listening comprehension, which describes three stages—perception, parsing and utilization. The theory holds that the second and the thirdstages(parsing and utilization) assist the first one—perception— through analyzing the grammar structure, the context, the usage patterns and conventions of a given utterance. Parsing and utilization are essential for the reconstruction of unclear speech sounds which arefound in real-life listening context. While the main condition for effective parsing and utilizing by ESL learners is their overall proficiency in English, their ability to use their existing knowledge of the language for the above mentioned processes might be improved through special training. The main idea of this study is that real-life listening practice may be beneficial and lead to faster improvement inESL students' listening comprehension. The method for this study involved repeated exposure of the students to noise-saturated recordings followed by multiple-choice tests. Also the students' performance on their regular ELC tests was analyzed in order to see if such practice was beneficial for the students in the treatment group. Statistical analysis of the study showed no significant difference that might be attributed to the training.

ACKNOWLEDGEMENTS

Table of Contents

Introduction

For a long time, a focus on listening in teaching and learning has been overshadowed by the other three ESL/EFL (English as a Second Language/English as a Foreign Language) skill "pillars"—reading, speaking and writing. Describing this pedagogical deficiency, Flowerdew and Miller (2012) noted, "Listening has often been described as the 'Cinderella skill"(p. 225). Since listening is a receptive skill (as opposed to the productive skills of writing and speaking), it has often been erroneously labeled as a passive skill. Such a characterization, of course, is inaccurate. Helgesen (2003) defined listening as "an active purposeful processing of making sense of what we hear (p. 24). Rost (2002) defined it as "the mental process of constructing meaning from spoken input" (p. 279).

Besides being underestimated, listening is arguably the most difficult language skill. ESL listeners have a very limited influence on the speech stream they hear and listen to. Except for the heavily adapted, beginner-level listening exercises, the process is a constant rush of the mind. As Henrichsen (1985) noted, "Another important difference between listening and all the other language skills—speaking, reading, and writing—is the amount of processing time available" (p. 53). Techniques like attending to body language, comprehension checks and confirmation checks can be used only by ESL learners in interpersonal listening. Further, in most cases, repeated comprehension checks will lower the native interlocutors' willingness to maintain communication. In some situations (e.g., during a lecture, or an exam, or in a movie theater) such checks are impossible.

ESL listening learners cannot consult a grammar book, choose to avoid complicated constructions, or look up unfamiliar words or expressions, which they can do when writing, speaking, or reading. Thus, if listening is "a mental process of constructing" (Rost 2002, p. 279), then the listener acts like a builder who must erect an edifice using materials he or she did not order. Listening in real-life circumstances often becomes even more complicated because of intruding background noise, other people's talk, and other distractions.

Teaching listening is not an easy task. As Goh (2014) wrote, "The challenges facing learners and teachers are due partly to a lack of understanding of what listening entails and how comprehension is achieved… Lessons typically involve listening to passages, answering questions, and checking answers" (p.72). Such activities test the students' ability to deal with information already taken in, but the listening comprehension as a dynamic multistage progression remains a mysterious act, and the process of turning the oral input into the aural intake seems to be an opaque thing-in-itself.

Brown (1987) observed, "Listening pedagogy was heavily influenced by practices in the teaching of reading, with little allowance made for the complex nature of the listening process and the unique features of spoken language" (p. 172). In teaching materials on reading and on listening we find many familiar techniques, such as inferencing, outlining, context building etc.

So, how can we approach the task of helping ESL students improve their real-life listening comprehension? First, we need to understand how comprehension is achieved when input is audible.

Different models and skills descriptions exist in connection with listening. For example, Vandergrift and Goh (2012) describe six main skills in the listening process: 1) Listening for details; 2) Listening selectively; 3) Listening for global understanding; 4) Listening for main ideas; 5) Listening and inferencing; 6) Listening and predicting.

However, for the purposes of this project, another model will be chosen—the classic one of Anderson (1980).The model posits three stages of listening comprehension—perception, parsing and utilization. Perception is a process in which the listeners compare the speech sounds they are listening to with semantic units in their mental lexicon. For example, analyzing the sounds of the word *bin*, the listener must discriminate the word-initial phone */b/* from its aspirated voiceless counterpart */p/*. The task may become more difficult if (1) the voicing is not heard clearly, (2) the alternative phone fits another meaningful word, like the second lexeme in a minimal pair, *pin*, or (3) the word cannot be analyzed grammatically and/or semantically within a larger context, like in the sentence *The password is 'bin'*.

Anderson(1980) defined the second stage, parsing, as "the process by which words in the message are transformed into a mental representation of the combined meaning of the words" (p. 402).For instance, the phrase *I have been*, being a high-frequency expression, is not likely to be misinterpreted even if the voicing of the phone *b* in the last word is weak to the point of sounding almost as the phone *p*. In this case the perceived alternative (*I have pin*) is filtered out as ungrammatical (an article *a* or *the* is missing).

"The third stage is the utilization stage, in which comprehenders actually use the mental representation of the sentence's meaning" (p. 402). For example, discriminating the phones in the aforementioned minimal pair *bin—pin*, even in less than ideal listening circumstances, may be assisted by the analysis of neighboring words, e.g., adjectives—*full bin* and *razor-sharp pin*. The pragmatic link between the adjectives and the nouns does not allow for alternating. *Full pin* or *razor-sharp bin* are senseless constructions.

Utilization processing may last longer than perception and parsing, and it might include analyzing a wider context at the discourse level. In the phrase from a piece of poetry, *Believe me when I beg you*, the word *beg* seems to be the best choice among several other alternatives, like *pack, peg, bag*, etc. However, utilization, in this case, completes its job only when the input is analyzed as a whole, with all the contextual elements comprehended. In the above-mentioned example, the verb *back*—which would make perfect sense in a different discourse—is not filtered out until analyzed against the whole background of the piece of poetry, the main message of which is romantic relations that lead to anguish and pleading.

J. R. Anderson sums up the explanation of the model in this way: "The three stages—perception, parsing and utilization—are by necessity partially ordered in time; however, they also partly overlap. Listeners make inferences from the first part of a sentence while they are perceiving a later part" (p. 402). I would also add that they are mutually beneficial. As was shown above, it is often impossible, during the perception phase, to discriminate between the minimal pairs without parsing

and/or utilization.

Along these same lines, Chomsky (1966) wrote, "There is good reason to believe that even the identification of the phonetic form of a sentence presupposes at least a partial syntactic analysis, so that the rules of the generative grammar may be brought into play even in the identifying of the signal. This view is opposed to the hypothesis that phonetic representation is determined by the signal completely, and that perceptual analysis proceeds from formal signals to interpretation" (p. 49).

This is a very insightful observation. The "syntactic analysis" or, under the time pressure of real-life listening, quick "syntactic matching" with the listener's mental lexis does not seem to start when the parsing phase is over. Perception—using the model of Anderson (1980)—cannot be executed in isolation from other phases of comprehension except for the rare cases when isolated words constitute the input.

When a native listener and an ESL student, both having approximately the same level of education and hearing acuity, listen to an authentic stream of English speech, the native listener has a substantial "top-down" advantage, his background knowledge of the language being incomparably richer. As a result, the ESL student will always be behind the native listener in processing the input. Moreover, native listeners may have the advantage over L2 learners even when the native listeners' hearing ability is impaired. I witnessed this myself when very old Americans didn't have problems with understanding a Hispanic speaker whose accent was very heavy but who was speaking loudly enough, while I, being a non-native listener with no hearing problems, was experiencing great difficulties trying to comprehend what the man was talking about. It seems then that the higher phases of listening comprehension may greatly facilitate the initial phase—perception.

The pedagogical question is this. How can such facilitation become more achievable for ESL learners? What are the ways of improving non-native listeners' perception by getting them to do something more than just matching their vocabulary storage with the sounds they are hearing and processing?

Most of the techniques used for teaching ESL listening are designed to help the learners cope with their listening tasks and exercises in the classroom or in the lab. However, while assisting students in doing their assignments, these techniques do not seem to give students the chance to develop their ability to process audible information in less than ideal, authentic situations of everyday life where, as a rule, no scaffolding is provided and noise and other distortions disrupt the audio signal. In the absence of such training, we naturally see a certain drop in ESL learners' listening skills outside the classroom.

Are there special ways of improving L2 learners' listening comprehension of authentic speech, particularly during the perception phase? Can practical instruction be given that exposes ESL learners to harder-condition listening tasks in order to prompt them to do their best, to raise their motivation, to help them leave their comfort zone, and to improve and help them become better prepared for genuine listening and interacting situations?

Literature Review

In this section, three key studies that deal with listening comprehension of both native and non-native listeners in less than ideal circumstances will be analyzed.

> In everyday life, the conditions for recognizing speech can be difficult. Traffic, loud music, for instance, are all sources of noise, which may hinder speech recognition. Listening in the presence of noise is hard in one's native language but even more challenging when listening in a non-native language, and gets even more challenging for both native and non-native listeners with increasing levels of noise (Cumans, van Hout, and Scharenborg, 2014, p. 1).

With this concept guiding them, Cumans, van Hout, and Scharenborg conducted a series of

experiments on 47 Dutch students using ESL listening materials with different signal-to-noise ratios. Their findings showed:

> Non-native listeners are able to use both word-initial and word-final information for word recognition, but fewer words were recognized with increasing difficulty of the listening conditions when the onset of words was masked. Thus, word-initial information is more important than word-final information for word recognition when listening conditions become harder. This increasing effect occurred independently from the proficiency level in the nonnative language of the participants, although proficiency level was correlated to test performance in general (p. 1).

In his article "Sandhi-Variation: A Filter of Input for Learners of ESL," Henrichsen (1984) hypothesized that the quality of one's perception depended not only on the quality of the signal but also on one's knowledge of English. He wrote, "Since listeners with a greater knowledge of the language system are less dependent on the signal itself, they will have little difficulty converting input to intake when clarity of the signal is reduced. Conversely, because listeners with less knowledge of the code are more signal dependent, they will have greater difficulty comprehending input when the perceptual saliency of its forms is reduced by a process such as sandhi-variation" (p. 108). Sandhi-variations are phonological modifications of morphemes when they are juxtaposed with other morphemes. Such modifications are frequent in authentic speech.

The following diagram shows the expected impact of sandhi-variations on the performance of native listeners, high-level ESL learners and low-level ESL learners.
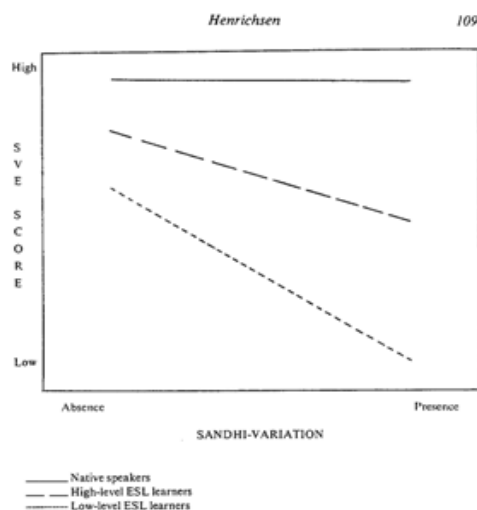
Figure 1. **Hypothesized interactions of presence/absence of sandhi-variation and level of language proficiency. From Henrichsen, L.E., (1984) Sandhi Variation: Filter of Input for Learners of ESL. In** *Language Learning*, **ed. Larsen-Freeman, D., vol. 34, no.3**

The actual results of the experiment showed a significant interaction between level of proficiency and how the participants scored dealing with exercises with and without sandhi-variation. The native speakers seemed to be practically unaffected by the differences in clarity of the input signal because they had an adequate knowledge of English that compensated for the reduced saliency of the input (p. 117). The situation with language learners was quite different.

> On the other hand, language learners with only a limited knowledge of the code are not able to compensate for the missing signal information when input is characterized by sandhi-variation. Consequently, they experience difficulty comprehending it. For even the high-level ESL learners, there is a significant difference between their presence and absence scores. When the clarity of the input signal is high (i.e., sandhi-variation is absent), their comprehension is at the level close to that of the native speakers and significantly different from that of the low-level ESL learners. However, when sandhi-variation is introduced the high-level ESL learners' scores drop dramatically. Thus, in the presence condition, there are significant differences between the mean scores of the high-level learners and the native speakers, but not between those of the high-level and

low-level learners (pp.118-119).
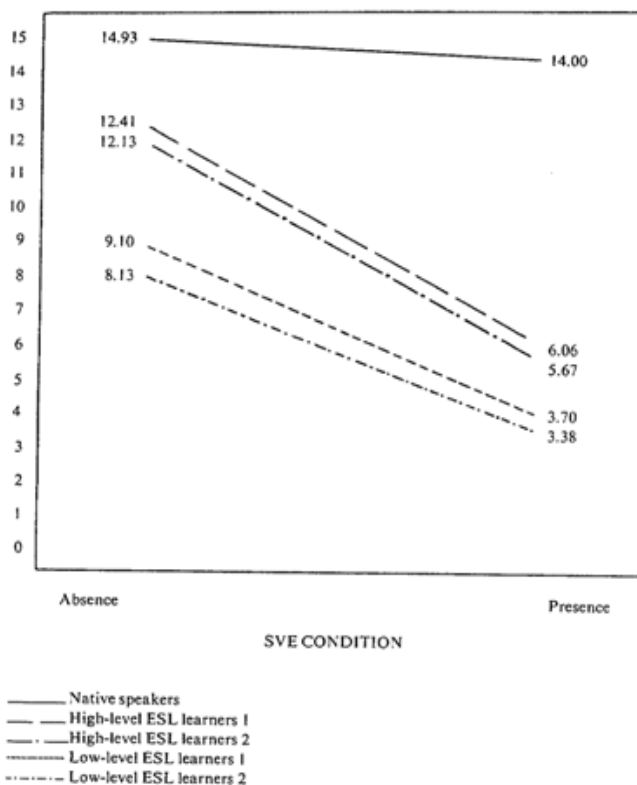

Figure 2 illustrates this situation:

Figure 2. Obtained interactions of SVE scores in presence and
absence conditions and level of subjects' language proficiency.
From Henrichsen, L.E., (1984) Sandhi Variation: Filter of Input for
Learners of ESL. In *Language Learning*, ed. Larsen-Freeman, D., vol.
34, no.3

It is interesting to compare the actual results with the experimenter's expectations. As we see,

the drop in comprehension for the high-level ESL learners was worse than hypothesized. As

Henrichsen noted earlier in this article, the native speakers seemed to be practically unaffected by the

differences in clarity of the input signal because they had an adequate knowledge of English, which

compensated for the reduced saliency of the input. It would seem logical to assume that both higher-

and lower-level ESL learners had to use the same, though less effective, compensatory mechanism to

deal with less-than-ideal-condition listening. The learners might have done better, had they been

trained in the ways of activating their background knowledge, of utilizing what they already had.

Other researchers have focused on other factors with listening. For example, Lecumberri, Cooke, and Cutler (2014) listed several of less-than-ideal-condition listening factors: channel distortion, reverberation, additive noise etc. (p. 871). The authors add,

> Another factor which differentiates L1 from later language acquisition is the quality and diversity of the input, ranging from single-source, non-native, heavily L1-accented pronunciations of the TL [target language] found in some FL [foreign language] learning situations to diverse, native and variable speech characteristic of natural contexts. Training studies have shown that for the formation of robust categories, diverse native input is necessary… It is notable that there have been few studies of the effect of simulated adverse conditions on training (p. 870).

The last sentence and especially the last word of this statement are very important. Many publications exist that describe and/or create classifications of adverse conditions and discuss the non-native listeners' performance in the presence vs. in the absence of these conditions, but little is found about the potentially beneficial impact of such conditions on non-native listeners' training.

The idea of not making the ESL learner's life too easy seems very logical. For example, at least one theory supports the beneficial effects of making the learners' tasks reasonably more difficult. Krashen's input hypothesis (1987) states that "the acquirer should understand input language that contains structure *a bit beyond his or her current level*[emphasis added] of competence" (p.20). Krashen defined the learner's current stage of studying language as i. The next stage of developing language is $i+1$. We must note that even though this idea is presented as a formula, it is not a quantitative expression but rather a qualitative statement, because it is very difficult to establish and test one's $i+1$. The idea, however, is clear. The level of language which learners are exposed to should be just a bit beyond their current competence. This line of thinking leads to the question: will the same

principle work if the quality of the input which the learners are exposed to is less clear (i.e., more authentic)?

Real-life-condition listening training may become a valuable addition to the existing set of instructional practices. It may act not as a rival of scaffolding but as a complement. With this hope in mind the current study was conducted—in order to test the practical effects of the real-life listening exercises.

The Research Question

The research question of this study was the following. *Will listening exercises in real-life-listening conditions affect beneficially the listening comprehension of the treatment group of the Advanced Low ACTFL level ESL participants?* The study was conducted at the English Language Center of BYU for the space of fourteen weeks, and the effect of real-life listening exercises was to be shown by comparing the treatment group's summative performance on their ELC listening tests in the end of semester and the control group's summative performance on the same tests. Also, a short survey was used to determine what the participants thought about such exercises—whether or not they felt such practice was helpful to them or not, which was the qualitative part of the research.

Method

Setting

The study was conducted at the English Language Center (ELC) of Brigham Young University. This school offers intensive English as a Second Language courses for learners who come from many countries to improve their English before they can begin studying in American colleges and universities. The majority arrive from Latin America. Also, many students come from Asian countries, such as South Korea, Taiwan, China, Thailand, and Japan. Some students come from Africa and

Europe. The ELC employs well-trained staff, professionally developed curricula and methods, and state-of-the-art technologies.

The program is designed to provide a laboratory school for teachers learning to teach English to speakers of other languages and to provide high quality English language instruction for students desiring to improve their English language skills. Students can expect most of the teachers at the ELC to be graduate students studying in the TESOL master's degree program or to have graduated from that program with a graduate certificate or master's degree.

At the ELC, language instruction and testing take place Monday through Friday from 8:00 a.m. to 3:00 p.m. Students are expected to complete 2 to 3 hours of homework assignments daily. They also participate in social activities sponsored by the ELC. Above all, the intensive nature of the program supports language development; students attend 20 hours a week of in-class instruction plus lab work (ELC, 2015).

Participants

Our study involved 33 non-native listeners placed at the Academic B level at the English Language Center of Brigham Young University, which corresponds to the Advanced Low ACTFL level. The participants suitability for the above mentioned level was established by pre-semester placement testing. The participants belonged to pre-existing groups based on class assignments (17 participants in the treatment group and 16 participants in the control group). Both groups were taught the Academic B Listening and Speaking Course by the same teacher. The treatment group met at 8:15 a.m., and the control group met at 9:30 a.m. four times a week, from Monday to Thursday. The treatment and the control groups studied in similar classrooms.

All the IRB requirements were met. Data were collected under the auspices of the ELC's permission to conduct research as a laboratory school.

Treatment

The participants in the treatment group were exposed to listening in real-life conditions, which was expected to encourage the learners to do their best, to raise their motivation by moving them out of their comfort zone and thus become better prepared for authentic interaction situations. In order to create real-life listening materials, we had to select and use a distractor.

Distractors are factors that lower the ease of perception and force the listeners to strain. Among these factors are high-paced speech, specially modulated speech, two parallel overlapping speech streams, speech with varying degrees of background noise, visual distractors, heavily accented speech, etc. Based on the perceived practicality of the experiment, the decision was made to use white unmodulated noise. This kind of noise is called *white* because the strength of the noise signal is spread evenly over the whole range of audible frequencies. The term *unmodulated* means that the amplitude of the noise signal is not affected by changes in the amplitude of the voice signal. The noise remains the same regardless of the speakers' voices sounding softer or louder. The voice signal was 50% (6dB) louder than the noise signal. This ratio (+6dB) is an established convention in the domain of acoustics.

> Physically doubling the sound pressure means to increase the sound pressure level by 6dB… On the other hand, if you reduce the sound pressure by 50% (e.g. from 1 to 0.5 Pa) the level has been reduced by 6 dB. (Norsonic, 2016)

The participants from the control group listened to clean recordings containing authentic and near-authentic discussions taken from a validated, level-appropriate text book, *Quest 3* (Blass,Hartmann,2007), which is designed for Advance Low ACTFL level ESL learners and sold commercially;

The treatment group listened to the same recordings mixed with the specially measured white noise. Every group had 11 listening sessions, spending 10-15 minutes at the end of a class period once

a week. At the end of each session, they took a brief multiple choice quiz to assess their comprehension of the input. The study was conducted for the duration of 11 weeks of a standard 14-week semester.

With the help of a computer program (Audacity), the original recordings were conditioned to approximate real-life listening. As was explained above, this was done by mixing the voice signal with a white unmodulated noise signal. Following is an explanation of the process we used to condition sound files.

Step 1.

Sound files (mp3) obtained from the Quest 3 Audio CD with the help of a sound-converting software (Free Audio CD to MP3 Converter 1.3.11) were opened in Audacity®, as shown in Figure 3.
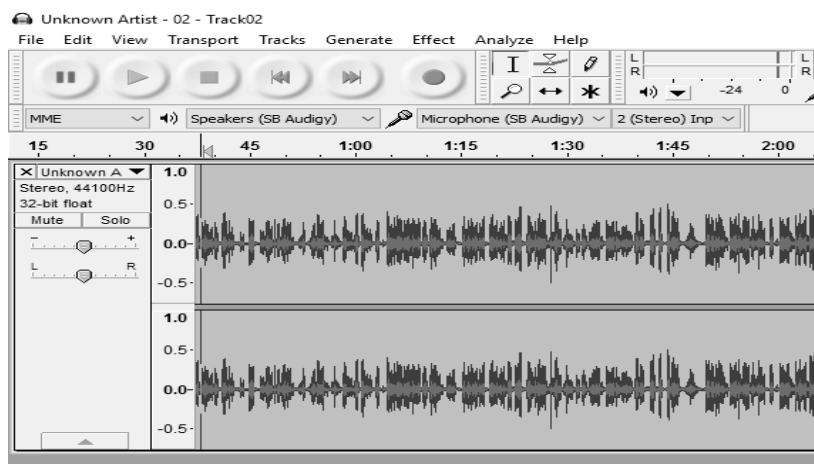

Figure 3. Opening mp3 files in Audacity

Step 2.

As shown in Figure 4, an additional blank track was created for each of the opened mp3 files.
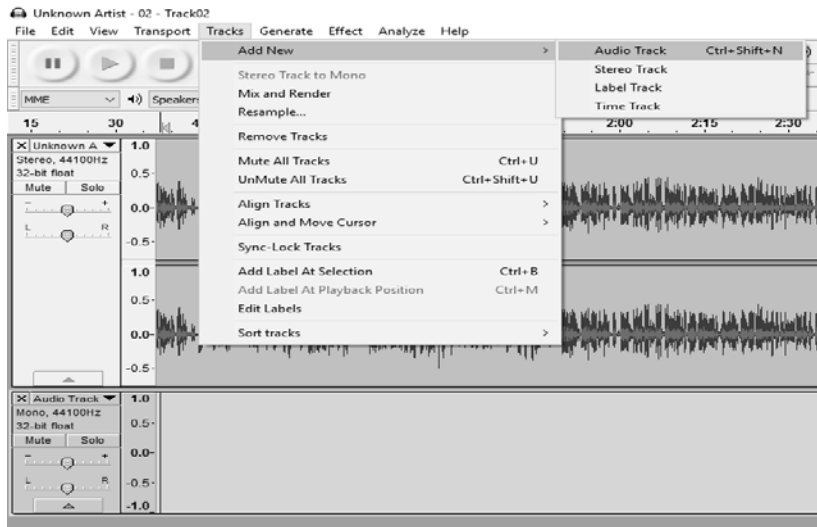

Figure 4. Creating an additional blank track

Step 3.

As shown in Figure 5 and Figure 6, the mean amplitudes of the sound tracks were found, with the help of Audacity's plugin "Wave Stats." The parameter presented by the plug-in is RMS, which stands for "root mean square." It is also known as the quadratic mean, defined as the square root of the arithmetic mean of the squares of a set of numbers. Root mean square (rather than a simple mean) is used because of the dynamic nature of sound waves.

> The root-mean-square pressure (abbreviated as RMS pressure) is the square root of the average of the square of the pressure of the sound signal over a given duration. The root-mean-square pressure is most often used to characterize a sound wave because it is directly related to the energy carried by the sound wave, which is called the intensity. The intensity of a sound wave is the average amount of energy transmitted per unit of time through a unit area in a specified direction. The intensity of a sound wave is proportional to the average over time of the square of its pressure. (Discovery of Sounds in the Sea, 2016)
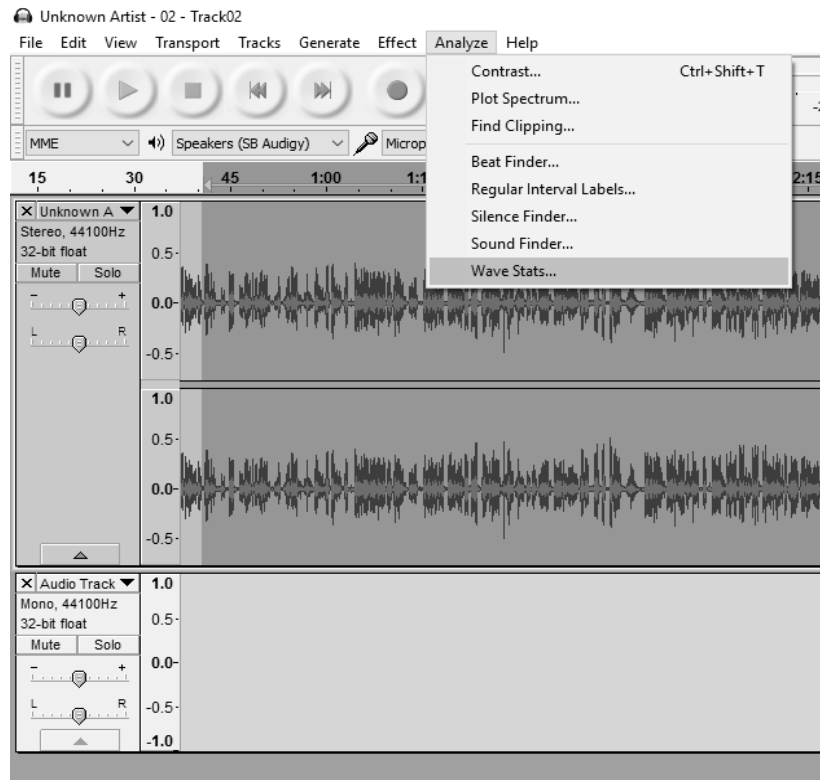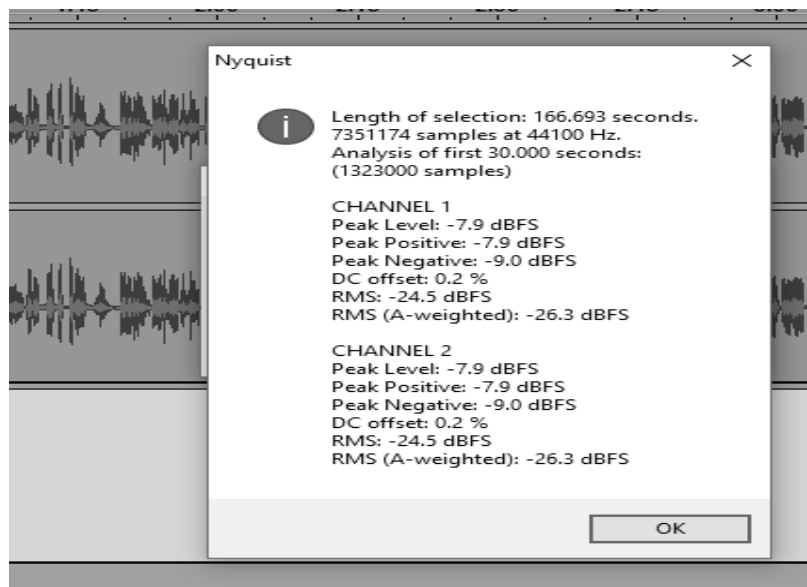
Figure 5. Using Wave Stats plug-in

Figure 6. Obtaining Root Mean Square

Step 4.

As shown in Figure 7, White unmodulated noise was generated for the additional track that had

been created earlier. The parameters of the noise amplitude could range from 0 to 1 (which is from 0%

to 100%). Since in the domain of Acoustics the 6 dB is universally accepted as a subjective doubling

the loudness, we used this stepping for generating noise. Setting the noise amplitude to 0.05 (or 5%) resulted in -6 dBFS signal/noise ratio, which meant, the speech signal was 6db louder than the noise.[i]
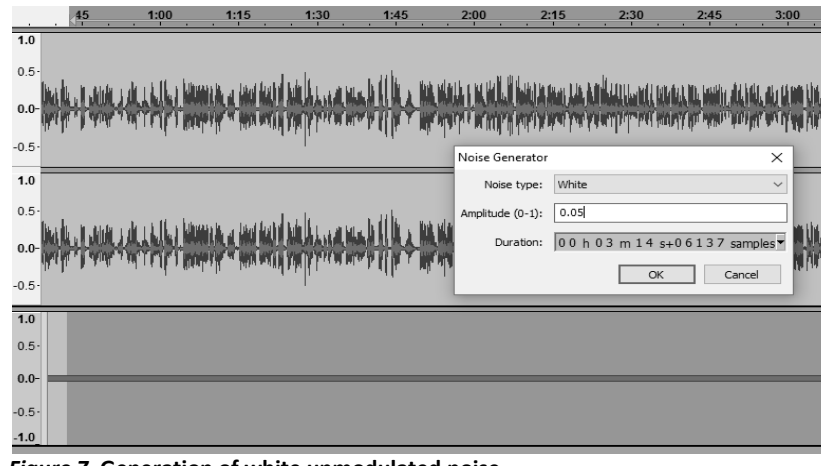
Figure 7. Generation of white unmodulated noise

Step 5.

The needed selection was saved as a conditioned file by using the "Export" command from the File menu.In the process of exporting the needed selection, an additional dialog window appeared, in which we chose the "Joint Stereo" option which allowed the two signal tracks to merge with the noise track which was added equally to both signal tracks, as shown in Figure 8.
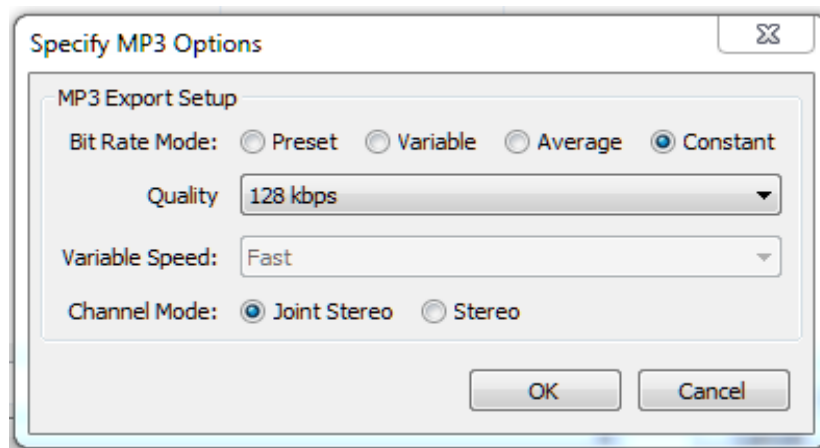

Figure 8. Merging the channels

The equipment for the study consisted of the following:

A) LG v10 smartphone as the playback device.

B) GPX 2.1 Portable speaker system with a built-in amplifier. Model NO: HT128.

C) Converting software: Free Audio CD to MP3 Converter 1.3.11 (freeware)

D) Conditioning software: Audacity ® 2.0.0 Unicode (freeware)

E) Playback software: Poweramp for Android, version 2.0.10-build-588-play (licensed

software)

F) Sound files format: MP3. Discretization frequency: 44.1 kHz; bit rate: 128 bps.

Instruments

The choice of testing instrument was made with the consideration of the necessity of using validated materials. The participants' listening comprehension in the study-related exercises was assessed weekly based on multiple choice quizzes taken from the *Quest 3*teacher's manual (Blass,Hartmann,2007).

The participants in both groups had taken a series of regular tests administered by their ELC teacher, which was a demand of their course syllabus. The students' listening comprehension was assessed through validated multiple choice quizzes, and the descriptive statistics obtained in the course of the study is presented in the "Analysis" section below. The inferential statistical data obtained from their ELC listening tests were assessed in order to find out whether or not the real-life-condition listening affected beneficially ESL students of the above-mentioned level (Low Advanced ACTFL). The data that reflected the participants' summative performance came from their ELC listening tests.

Results and Discussion

The results obtained from study-related evaluation are presented here as descriptive statistical data in the form of tables and histograms. Also, a qualitative analysis of participants' responses to a

short, six-question survey is included. However, the main point of interest was the results of participants' ELC listening tests taken during the semester. This analysis was made through a Repeated Measures ANOVA test.

Analysis of Data

Descriptive quantitative data from the listening exercises

The real-life-condition listening exercises conducted in the treatment group and in the control group yielded the following quantitative data.

The mean performance in the treatment group was 69.34%.The mean performance in the control group was 70.27%.

Figure 9 shows the results of study-related multiple-choice quizzes taken by both the treatment and the control groups. Although the scores varied due to variations in the difficulty of the quizzes, the control group's performance is consistently higher than that of the treatment group. The similar configuration of the two lines gives us the reason to conclude that the white noise present in the recordings the treatment group listened to was the cause of this difference.
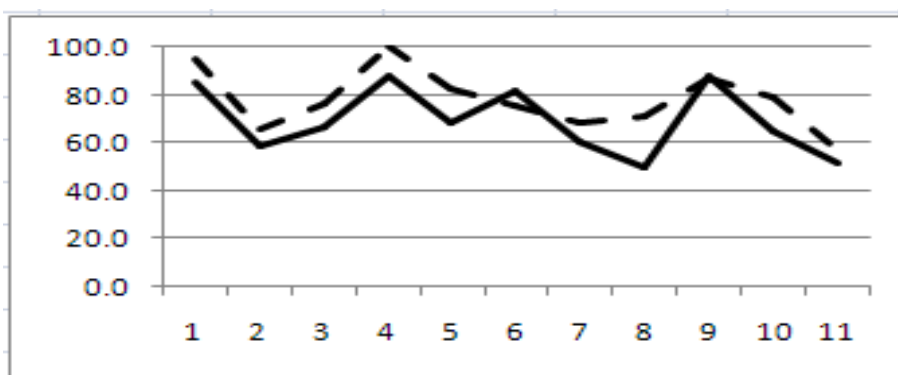

Figure 9. Performance on the study-related tests
in both groups.
———— The solid line is the treatment group
- - - - - - The dashed line is the control group

Qualitative assessment of the study-related exercises

A qualitative assessment of students' attitudes was also conducted in the treatment group. At the end of the study, the participants were emailedquestionnaires with the following questions about their opinions and attitudes in connection with the real-life-condition listening exercises.

a)   Have you been comfortable about listening to the noise-saturated recordings or has it been irritable?

b)   Do you think such practice can help foreigners to faster improve in their comprehension of spoken English?

c)   Do you feel the real-life-condition listening exercises have helped you improve your listening comprehension?

d)   Do you think the noise on the recordings was too loud?

e)   Do you think the noise on the recordings was too low?

f)   What would you suggest should be changed as far as the content of the recordings?

Because of the timing and fact that it did not affect their grades, only three participants out of seventeen returned their surveys, which is about 18%. Others failed to do so. The responses of the participants (T02, T05, and T11) are presented after each of the questions.

a)   Have you been comfortable about listening to the noise-saturated recordings or has it been irritable?

T02: I have been comfortable.

T05: The first time that we had the test, I feel irritable and frustrated a little bit, but then I got easier.

T11: It was like doing the test or reading with the many voices on the background, which I getused to. I kept my focus, and the noise didn't really interrupt me.

The responses to this question show no consistently negative reaction to the real-life-condition listening exercises.

b)   Do you think such practice can help foreigners to improve in their comprehension of spoken English faster?

T02: Yes, I do.

T05: We never listen to any language in complete silence. Even on tests, like TOEFL, we have noise. I think this kind of exercise can prepare students to listen to English in the most adverse-conditions of noise and also get used to the noise.

T11: Maybe. I'm not sure for the speaking part - that we will speak faster or not. But for the listening, I think this will help a lot for us to practice to listen and focus on what people on the test record are talking rather than of the background voices.

The answers ranged from solid positive to cautious neutral. The question was aimed at eliciting the participants' subjective opinion, and even with only three respondents a range of opinions was represented.

c)   Do you feel the real-life-condition listening exercises have helped you improve your listening comprehension?

T02: Yes, I do. I feel I can now be really focused on the conversation or lecture.

T05: Surely, yes!

T11: Absolutely yes! Actually today I just had the LATs test at the school, and there were a lot of people speaking and answering questions through their microphones. Before this semester, I tried not to listen to them, but my mind still listened to them a little bit. But for today, I didn't listen to any of them at all, so I really focused on my test and I for sure felt that I would get higher score this time. So, thank you so much!

Even though (as will be shown hereinafter) no significant quantifiable difference was shown in the treatment group's performance on regular (i.e., noise-free) tests, the three participants expressed a clearlypositive attitude towards their personal real-life-condition listening experience.

d)   Do you think the noise on the recordings was too loud?

T02: No, I don't. It was loud enough.

T05: It was OK.

T011: First time yes, but later on - not at all ☺

e)   Do you think the noise on the recordings was not loud enough?

T02: No, I don't. It was loud enough.

T05: It was loud enough.

T11: Sincerely, I don't know because I didn't pay attention to them at all. Sorry for that!

Questions "d" and "e" dealt with the technical parameters, which seemed to be agreeable enough for the participants.

f)   What would you suggest could be changed as far as the content of the recordings?

T02: I think everything was good.

T03: Hard to say.

T11: I'd like to suggest experiments with noise made by human conversations as a background sound, simulating a real event.

Participant T11 expressed an interesting idea about using a distractor that in real life is present no less often than noise. However, using it for gathering quantifiable data might be even more complicated, since this sort of background sound cannot begenerated by software, which raises the problem of proper measuringthe sound level of such background talk and mixing it accurately with a voice signal.

Had the response rate been higher, it would be possible to state that the students' attitudes were generally positive. However, it is possible that those students whose attitude was negative simply ignored the survey.

ELC test results and analysis

The two groups' performance on their EL:C tests was of the greatest importance for this study. The working hypothesis was that participants in the treatment group would demonstrate a significant positive difference compared to the control group on the regular listening tests administered by their ELC teacher during the semester.

To establish the acceptable parity of the treatment and the control groups at the onset, the pre-semester assessment of the two groups, scores on the pre-semester test were analyzed, and the results were following:

The treatment group: mean—71.51%; standard deviation—53.61.

The control group: mean—72.71%; standard deviation—64.61.

Also, the students' scores on the ten listening tests administered during the semester were analyzed. Table 4 in the Appendix section shows the participants' performance on those tests, as well as the results of the pre-semester test.

The mean of the ten regular ELC listening tests both groups had had during the semester was 73.80% for the control group, and 72.50% for the treatment group.

The assessment procedure used to determine if any significant difference existed between the two groups was the Repeated Measures ANOVA. Table 1 shows the results of this measurement.

Table 1
Repeated measures ANOVA results for the pre-semester and post-semester results in the control group and the treatment group

	Pretest			Posttest		
Group	M	SD	N	M	SD	N
Control	727.067	64.608	15	737.733	58.459	15
Treatment	715.125	53.610	16	634.438	249.704	16
Total	720.903	58.496	31	684.419	188.479	31

The results of the Repeated Measures ANOVA did not show a significant difference between the treatment and the control groups: $p=.185$; $\eta_p^2 = .06$. The null hypothesis is not rejected when p-value is larger than .05, and we have .185. While the partial eta squared (η_p^2) does not directly define the conclusion, it serves as a suggestion that is used to either accept or reject the working hypothesis. The value of .06 lies between the small (.02) and medium (.13) probability of the truthfulness of the working hypothesis. Thus, we concluded that real-life-condition listening practice did not lead to improvement in students' listening comprehension, under the conditions of this study.

Limitations of the Study and Recommendations for Future Research

The study was an attempt to detect if there were an observable positive effect from real-life-condition listening exercises in two Advanced Low ACTFL level groups of students. No observable positive effect was found. So, based on the results of this study, we may assume that adverse-condition listening practice may not help students improve their listening accuracy. However, this study does not seem to allow us to make such a decisive conclusion. It had many limitations. Among limiting factors that might have contributed to this result we can list the following.

1. The short duration of the study. The study lasted for only eleven weeks, due to the time frame of the semester (fourteen weeks).

2. The low frequency of the real-life-condition listening practice. The exercises were administered only once week. More frequent exercises may yield a different result.

3. The short duration of the exercises' sessions. The sessions lasted for about 10 minutes only, which included listening to recordings for up to five minutes and taking a multiple-choice test, which lasted for up to 5 minutes.

4. The participants' insufficient English language proficiency. Parsing-assisted and utilization-assisted perception could have been inefficient due to the Advanced Low ACTFL level of the students. Hypothetically, participants at the AdvancedMiddle or Advanced High ACTFL levels would have shown a better ability to progress in their listening comprehension due to a higher level of their mastery of grammar, usage, as well as their larger internalized lexicon.

5. Insufficient motivation of the participants. Since this study was not a part of the students' normal curriculum, they did not feel it important enough for their academic success and/or standing. For example, only 3 of the participants returned their answers to a short survey that was emailed to them in the end of semester, which weakened the qualitative part of the assessment.

6. Simplified experimental and assessment design. Because of the limited time and manageability of the participants, the two groups were not both tested on listening comprehension in real-life conditions.

7. A possible timing variable. The control group had the benefit of the teacher's practice in the preceding hour. Therefore, they might have received better instruction.

8. Qualitative assessment (survey) was conducted after the last study-related listening session. Because of the low motivation of the students, only three of them return their answers. I would have been better if the survey were sent to the students earlier.

Future research along these lines would benefit from a longer study period with longer, more frequent real-life practice sessions. Another possible improvement (suggested by a participant in the survey) would be to use real, human background noise instead of white, unmodulated noise.

Conclusion

Contrary to the expectations, the study did not show an observable positive effect of the real-life-condition exercises. As explained in the "Limitations of the Study" section, a lot of factors could have led to this result. It might be possible that real-life (i.e., adverse-condition) listening exercises proves to be a good way of training foreign L2 listeners and helping them achieve a better level of listening comprehension. However, that will have to be proven by future studies that do not suffer from the limitations of this study.

Taking into account the results, the limitations and the recommendation of this study, future researchers—with better conditions, resources, and a longer time for real-life-condition listening exercises—may observe a positive effect ofreal-life listening practice and the L2 listeners' improvement in their listening comprehension.

References

Anderson, J. R. (1980). *Cognitive psychology and its implications.* (4ᵗʰed.) San Francisco: W. H. Freeman and Company.

Blass, L., Hartmann, P. (2007) *Quest 3 Listening and Speaking*, (2nd ed.) New York, The McGraw-Hill Companies, Inc.

Brown, G. (1987). Twenty-five years of teaching listening comprehension. *English Teaching Forum. 25*, 11-15.

Chomsky, N. (1966). Linguistic theory. In R.G. Mead (Ed.),*Proceedings of Northeast Conference on the Teaching of Foreign Languages*(pp. 43-49). Menasha, WI: George Banta.

Cumans, L., Van Hout., Scharenborg, O. (2013) *Non-native Word Recognition in Noise: The Role of Word-initial and Word-final Informatio*n; Centre for Language Studies, Radboud University Nijmegen, The Netherlands.

Discovery of Sounds in the Sea (2016). Retrieved from http://www.dosits.org/science/advancedtopics/signallevels/

ELC Program Description (2015). Retrieved from http://elc.byu.edu/prospstudent/brochure/brochure2.php

ELC Levels (2015). Retrieved from http://www.elc.byu.edu/curriculum/levels.php

Flowerdew. J., Miller, L., (2012) Assessing listening. In C. Coombe, P. Davidson, B. O'Sullivan, S. Stoynoff (Eds.),*The Cambridge Guide to Second Language Assessment* (p. 225). Cambridge University Press

Goh, C.C.M., (2014). Second language listening comprehension: Process and pedagogy. In M. Celce-Murcia, D. Brinton, M. Snow (Eds.), *Teaching English as a Second or Foreign Language* (pp. 72-89). Boston: HeinleCengage Learning.

Helgesen, M. (2003) Listening. In D. Nunan (Ed.), *Practical English language teaching*. New York, NY: McGraw-Hill.

Henrichsen, L.E. (1985). Listening comprehension in C. C. Fries' Oral Approach. In P. H. Fries (Ed.), *Toward an understanding of language: Charles Carpenter Fries in perspective*. Amsterdam/Philadelphia: John Benjamins.

Henrichsen, L.E., (1984) Sandhi variation: A filter of input for learners of ESL. *Language Learning*, 34, (3)

Krashen, S.D., (1987). *Principles and practice in second language acquisition.*Hemel Hempstead: Prentice-Hall International.

Lecumberri, M.L.G., Cooke, M., Cutler, A. (2014).*Non-native speech perception in adverse conditions: A review*. Language and Speech Laboratory, Facultad de Letras, Universidad delPais Vasco, Vitoria, Spain. Retrieved from http://laslab.org/upload/non-native_speech_perception_in_adverse_conditions:_a_review.pdf

Krashen, S.D., (1987). *Principles and practice in second language acquisition.*Hemel Hempstead: Prentice-Hall International.

Norsonic. (2016). The Lever of Sound (dB). Retrieved from http://www.norsonic.com/en/applications/measuring_sound/the_level_of_sound_db/

Rost, M., 2002. *Teaching and researching listening*. Harlow, UK: Pearson Education/Longman

Vandergrift, L. &Goh, C. (2012) *Teaching and learning second language listening: Metacognition in action.* New York: Routledge.