

Comparing Authentic and Scripted Language Listening Comprehension in University-Level EFL Learners

Robert Long

Kyushu Institute of Technology

This study investigated whether Japanese university EFL learners' listening comprehension differed significantly on tests of authentic language (AL) or scripted language (SL). Possible differences in comprehension between low and intermediate proficiency students, and how lexical density of the listening samples correlated with listening comprehension, were also explored. Participants scored higher on SL multiple choice tests than on AL tests on average. There was no significant difference in the performance between low and intermediate proficiency students in this study, but there was a significant difference between scores of learners in the two pre-intermediate groups. The inclusion of more content words also appeared to aid the lowest performing students in better understanding authentic listening passages. The researcher suggests explanations for these findings, and suggests the need for more authentic listening practice in language classrooms.

本論は日本人EFL大学生のリスニング力が、オーセンティック・ランゲージ (AL) のテストあるいはスクリプト・ランゲージ (SL) のテストにおいて大きく異なるかどうかを調査したものである。初・中級者間における理解力の相違の可能性や、リスニング・サンプルの語彙密度とリスニング力の関係も調査した。SL多肢選択テストの得点の方がALテストよりも平均して高かった。初・中級者間には有意差はなかったが、2つの準中級学習者群間の得点には有意な相違が認められた。また、内容語をより多く含めることは英語力の低い学習者がオーセンティックなリスニング題材を理解する助けになっているようであった。これらの研究結果に対する説明と、よりオーセンティックなリスニング練習の必要性についても述べる。

With continued emphasis placed on students in many institutions in Japan to achieve high scores on standardized tests such as the Test of English for International Communication (TOEIC) and the *Jitsuyō Eigo Ginō Kentei* (EIKEN), educators are faced with the choice of using materials that resemble test listening passages or materials that use unscripted, authentic language (AL). However, such materials might not be perceived as helpful for test-taking. AL material also might be viewed by some students as too difficult.

However, to successfully deal with English as it is spoken outside of the classroom, learners must become familiar with the linguistic and sociocultural characteristics of English as spoken by members of

different cultures. Standardized listening passages, on the other hand, tend to eliminate many features of authentic speech, such as fragmented speech, repetition, and rephrasing.

To better understand how authentic language listening passages compare with scripted language listening samples, this study investigated several issues. The first issue was whether university EFL learners' comprehension of SL and AL samples differed significantly. The second issue was whether proficiency (as indicated by TOEIC scores) affected listening comprehension on both SL and AL listening samples to similar degrees. The third issue was how lexical density indices of listening sample transcripts impacted students' listening comprehension of the two varieties of listening tests.

Literature Review

Hedge (2000) argued that the notion of authenticity was popularized with the communicative approach to language teaching in the 1970s. Seven interpretations of authenticity emerged (Carter & Nunan, 2001; Herod, 2002; Herrington & Oliver, 2000; Jacobson, Degener, & Purcell-Gates, 2003; Jordan, 1997; Nunan, 1988; Stubbs, 1996). For listening, these resulted in three basic concepts: AL refers to language that (a) is relevant to students' lives, (b) models real-world situations, and (c) is not produced for language teaching purposes. Martinez (as cited in Berardo, 2006) listed the following advantages and disadvantages of authentic language use:

Advantages:

- Students are exposed to real language, which reflects language variation.
- There is factual acquisition.
- Textbooks do not contain inaccurate language that is found in authentic speech.
- Authentic materials might motivate students.
- One piece of text can be used for various activities.

- There is a wide choice of styles, genres, and formality in authentic texts.
- Authentic reading materials can make students eager to read for pleasure.

Disadvantages:

- Authentic language might be difficult to understand because of a culture gap.
- The vocabulary might not be immediately useful for learners.
- Authentic language might be too difficult for beginners.
- Preparation of the texts and activities is often demanding and time consuming.
- Accents and dialects can vary.
- Some materials (e.g., news broadcasts) might become outdated quickly.

For many researchers, authenticity is important because it prepares learners for realistic situations. According to Brown and Eskenzai (2004), by using textbooks alone, learners are not exposed to language as it is used in the real world. Using fewer authentic materials with learners might lead to problems in interactive competency. Otte (2006) argues that learners need to “practice using authentic language themselves, in order to be better prepared to deal with authentic language in the real world” (p. 56).

Liu (2016) found that the exploitation of authentic materials can be used as a bridge for students to better understand original sources. He wrote that the proper adaptation of various AL samples is needed to make them accessible to learners and that teachers should be aware of the strengths and weaknesses of the materials that are chosen. Furthermore, Yin (2015) explored how listening comprehension is improved by authentic listening practice. She examined the listening practices of 22 American ESL university students in an intensive English program. She conducted a statistical analysis to examine the relationship between the learners’ outside-of-class language activities and their listening comprehension performance in listening tests. Her results showed a significant correlation between authentic listening activities (such as communicating with native speakers and watching English television shows) and listening comprehension performance. Yin’s analysis also revealed a positive relationship between learners’ self-efficacy and their listening comprehension ability. In short, such research reinforces the idea that authentic language listening practice has a role to play in EFL learning.

Research Questions

Three research questions were investigated:

1. In this study, would there be significant differences between listening comprehension scores of authentic language and scripted language listening samples?
2. Do low and intermediate proficiency learners perform differently on authentic and scripted language listening comprehension tests?
3. Is there a correlation between correct answers and lexical textural density measurements of listening scripts?

Methodology

Participants

The 82 participants in this study were 18- and 19-year-old first-year students at a national university. The participants were engineering majors and came from three English communication classes that were held weekly. These classes focused on developing learners’ English speaking and listening skills. One class included 29 intermediate students (TOEIC range 450-600), and the other two classes together were composed of 53 low to low-intermediate proficiency students (TOEIC range 285-449). There were 13 female participants and 69 male participants.

Material

Eleven listening samples from GoldFish (www.goldfish365.com) and nine listening samples from Voice of America (VOA) Easy Listening (www.learningenglish.voanews.com) were used for this study, both of which are free online EFL resources. GoldFish is a teacher-produced English listening and fluency practice website that provides recordings of unscripted conversations. VOA Easy Listening provides scripted listening samples and texts about an array of topics, such as animals, health, and history. Each recording was four-to-six minutes long.

The recordings varied by vocabulary use and rate of speech, but were not sorted by difficulty levels on their websites. They were comparable only in that each recording lasted between four-to-six minutes. Quizzes for both authentic and scripted listening samples were created by the researcher and included 15 multiple-choice referential questions about details and key terms. These quizzes were read by fellow teachers and were judged to be appropriate for the participants of this study.

Procedures

Two listening tests were conducted in each class from July 2017 through February 2018, except for the months of August and September. Students listened to each sample once and answered 15-question comprehension quizzes. Each procedure lasted about 10 minutes. *BBC News Report* (BBC, 2015) samples were also used occasionally, but students' scores for these quizzes were not collected. Twenty quizzes were given in total, with 11 from GoldFish website and nine from VOA. Only data from those quizzes were analyzed for this study.

Results

The statistics for each group's listening AL and SL comprehension scores are shown in Table 1. To answer the first research question concerning listening comprehension scores on SL and AL listening tasks, a paired-samples *t*-test was conducted, and assumptions for parametric statistics were checked and met. The results showed a significant difference between average scripted and authentic quiz scores, $t(81) = 12.18, p < .001$, with a large effect size ($d = 1.35$). This suggests that students found the AL quizzes significantly more difficult than SL quizzes.

Table 1. Participants' Average Quiz Scores by Group and Listening Type

Groups	Quiz Types	Min	Max	M	SD
Lower group 1 (<i>n</i> = 28)	AL Average Scores	6.8	10.6	8.8	1.2
	SL Average Scores	7.6	12.8	9.9	1.4
Lower group 2 (<i>n</i> = 25)	AL Average Scores	5.6	11.2	9.4	1.5
	SL Average Scores	7.5	12.9	11.3	1.0
Inter-mediate (<i>n</i> = 29)	AL Average Scores	6.0	11.8	9.0	1.4
	SL Average Scores	7.6	13.2	10.9	1.5

Note. AL = authentic listening, SL = scripted listening.

The second research question was whether there was a significant difference in comprehension scores between the different groups of learners. To answer this question, a repeated measures ANOVA was run. A statistically significant difference was found with a medium effect size, $F(2,79) = 4.591, p = .013, \omega^2 = 0.081$, so post hoc comparisons were

made using the Bonferroni correction to control for type 1 errors. No significant differences were found between the intermediate proficiency group's AL and SL quiz scores and those of either lower proficiency group. However, there was a significant difference between the scores of the two lower proficiency groups, with students in group two performing significantly better on the listening quizzes (see Figure 1). This surprising result is discussed later.

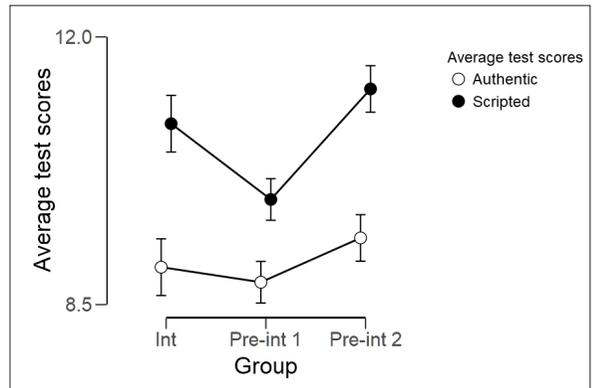


Figure 1. A descriptive plot of average listening scores per group.

Note: Error bars represent 95% confidence intervals.

The third research question asked whether a correlation existed between participants' scores and listening texts with higher or lower lexical density values. As shown in Table 2, the average lexical density was higher for the scripted than the authentic texts. An independent samples *t*-test showed that this difference was both significant and had a large effect, $t(18) = 4.59, p < .001, d = 2.06$.

Table 2. Average Lexical Density of the Authentic and Scripted Listening Texts

Text type	Min	Max	M	SD
Scripted	0.51	0.73	0.63	0.072
Authentic	0.38	0.58	0.49	0.060

Correlations were calculated between the average scores of all participants on all quizzes and the average lexical density of all texts. Both categories were also broken down by student group and text type. As Table 3 shows, there was a large and statistically significant overall correlation between the average quiz scores and the average lexical density of the texts, $r = .559, p = .01$. This correlation is positive,

which means that students performed better as lexical density increased, which seems counter-intuitive.

Although breaking down the results by student group and text type results in very small sample sizes, which must be interpreted with caution, this analysis does suggest a possible explanation. As can be seen in Table 3, the positive correlation between average quiz scores and average lexical density of the listening texts is driven by the authentic texts. The scripted listening texts show no significant correlation. In addition, the lower proficiency group one, the group with the lowest average quiz scores, is the only group which maintained a significant positive correlation between quiz scores and average lexical density of the authentic texts. This suggests that natural language features of the authentic texts, such as fillers or increased grammatical complexity, were particularly detrimental to the students' comprehension in this group. Implications of these findings are considered in the discussion.

Discussion

Comprehension scores between the low-intermediate proficiency EFL learners and the intermediate learners indicated that both levels of students found authentic listening more difficult to comprehend than scripted listening in this study. Although the data show that participants scored higher on the scripted listening tests, the scripted language recordings in this study often presented various social, political, and economic topics, and therefore used more academic language that likely should have been more difficult for the learners. Therefore, the lower AL test scores might have been a result of other features of authentic speech. It is possible that faster speaking rates, more common use of slang, fragmented utterances, repetitions, and rephrasing were factors that impeded comprehension though this should be investigated in further research. This suggests a need for more exposure to these features if EFL learners are to recognize them. Cruz (2018),

the website designer of www.goldfish365.com, also proposed that if authentic listening materials can be found which are similar to scripted materials in regards to topic, length, and difficulty, then AL tasks should be preferred because they help students both with test results and to develop real-life communication skills. Moreover, authentic materials help students to be more conscious of high frequency vocabulary that is used in daily life, and how native speakers articulate their words allowing them to become accustomed to the features of authentic speech.

Another feature that could have impacted comprehension was background knowledge. Participants might not have had the relevant background knowledge to adequately understand and synthesize many of the ideas and details that were presented in some of the recordings. Therefore, background knowledge, rather than readability, might have determined which recordings were more or less difficult.

As for the second research question concerning the performance of low and intermediate proficiency learners and how they performed in both kinds of listening comprehension quizzes, students in pre-intermediate group two did perform significantly better on the listening quizzes. The reason for this could be that the differences between the two classes were based on TOEIC scores and many of the students' scores were not far enough apart to show evidence of differences in proficiency. Furthermore, TOEIC scores take into account reading, so many students could score higher on this section and lower in listening and still obtain a relatively good overall TOEIC score. Initially the data showed that the intermediate learners and the pre-intermediate learners performed similarly. However, when the two pre-intermediate groups' scores were analyzed, it was found that one group scored higher, and one group scored lower, than the intermediate learners. It was assumed that the label of these intact groups would accurately represent differences in their listening abilities for this study.

Table 3. Pearson Correlations (*r*) Between Average Quiz Scores and Lexical Density of the Texts

Groups	All texts		SL texts		AL texts	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Lexical density - All scores	.559	.010	-.081	.836	.679	.022
Lexical density - Int scores	.550	.012	-.041	.917	.393	.232
Lexical density - Lower 1 scores	.390	.089	-.196	.613	.689	.019
Lexical density - Lower 2 scores	.553	.011	-.004	.991	.319	.340

Note. AL = authentic listening, SL = scripted listening.

These learners were grouped by TOEIC scores, but were not given a listening pre-test before the start of this study. This demonstrates how using intact groups for research purposes can potentially produce misleading results.

The third research question about correlations between students' scores and lexical density did show a significant and large effect, with average lexical density of the texts accounting for 31% of the variance in students' average quiz scores. As was stated before, the correlation was positive, so it indicates that more content words (nouns, verbs, adjectives, adverbs) aided student comprehension. The data shows that the group with the lowest TOEIC scores scored higher on quizzes for listening tasks containing many content words, which then made the recording more comprehensible for them. This is counter-intuitive, because past studies have shown that higher lexical density usually makes listening comprehension more difficult.

Limitations

One limitation of this study was that students within different proficiency groups had large differences in TOEIC scores. In order to obtain more meaningful data, educators should ideally compare two clearly different ranges of TOEIC scores (i.e., 400-500 and 650-750). While the factor of learners' background knowledge was not controlled for, other limitations relate to a lack of pre-testing, and lack of comparability in regard to lexical complexity and topic comparisons between the AL and SL texts. Furthermore, because the sample size was so limited, the correlation data must be taken with caution.

Conclusion

AL tasks serve to acquaint students with a variety of accents, speaking rates and styles, and idiomatic usages. They also help prepare students for realistic and open-ended interactions. Most importantly, AL exposes students to pragmatic features of natural English usages, such as how pausing is used to emphasize ideas, how turns are taken, and how particular ideas are emphasized. On the other hand, scripted language orients students to news-like speech, academic language, and more complex social and political topics. A combination of these two forms of aural input is necessary to help our learners comprehend all of the nuances of a second language.

Acknowledgments

This research was supported by a Grant-in-Aid for Scientific Research (KAKENHI) of the Ministry of Education, Culture, Sports, Science, and Technology in Japan (No. 15K02788). I also would like to acknowledge Mike Mackay for his help, and to thank José Cruz for his comments and for his technical help in website design for my website (www.genderfluency.com), and for developing <http://www.goldfish365.com>, which provided material for this study.

References

- BBC Easy Listening (2018). Retrieved from <http://www.bbc.co.uk/learningenglish/english/features/news-report>
- Berardo, S. (2006). The use of authentic materials in the teaching of reading. *The Reading Matrix: An International Online Journal*, 6(2), 60–69. Retrieved from <https://www.researchgate.net/publication/237413136> The Use of Authentic Materials in the Teaching of Reading
- Brown, J., & Eskenzai, M. (2004). Retrieval of authentic documents for reader-specific lexical practice. *Proceedings of InSTIL/ICALL Symposium 2004*. Retrieved from <http://www.cs.cmu.edu/~jonbrown/publications/InSTIL-2004.pdf>
- Carter, R. & Nunan, D. (2001). *The Cambridge guide to teaching English to speakers of other languages*. Cambridge, England: Cambridge University Press.
- Cruz, J. (2018, February 28). Personal interview.
- Jacobson, E., Degener, S., & Purcell-Gates, V. (2003). *Creating authentic materials and activities for the adult literacy classroom: A handbook for practitioners*. Retrieved from <http://www.ncsall.net/fileadmin/resources/teach/jacobson.pdf>
- Jordan, R. R. (1997). *English for Academic Purposes: A guide and resource for teachers*. Cambridge, England: Cambridge University Press.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. New York: Oxford University Press.
- Herod, L. (2002). *Adult learning from theory to practice*. Retrieved from http://en.copian.ca/library/learning/adult_learning/adult_learning.pdf
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology, Research, and Development*, 48(3), 23–48. doi:10.1007/BF02319856
- Liu, J. (2016). Adaptation of authentic materials in English listening comprehension classes. *Theory and Practice in Language Studies*, 6(9), 1774–1779. doi:10.17507/tpls.0609.08
- Nunan, D. (1988). *The learner-centered curriculum*. Cambridge, England: Cambridge University Press.

- Otte, J. (2006). *Real language to real people: A descriptive and exploratory case study of the outcomes of aural authentic texts on listening comprehension of adult ESL students enrolled in an advanced ESL listening course* (Doctoral dissertation). Retrieved from ProQuest Dissertation Abstracts International database. (UMI No. 3212979)
- StatCalc [Computer software]. (1999). Retrieved from <https://www.acastat.com/statcalc/statcalc.html>
- Stubbs, M. (1996). *Text and corpus analysis*. Oxford, England: Blackwell Publishers.
- Textalyser [Computer software]. (2004). Retrieved from <http://textalyser.net/>
- Yin, M. (2015). *The effect and importance of authentic language exposure in improving listening comprehension* (Master's thesis). Retrieved from http://repository.stcloudstate.edu/engl_etds/33/

Robert Long has a Master's degree in English Education and a Specialist Degree in Multilingual/Multicultural Education. He has worked at Kyushu Institute of Technology for 23 years, where he is currently a professor. He has written several EFL textbooks concerning oral communication and technical English. His current research focuses on fluency, dysfluency, authentic language, gendered and same-sex interactions, and writing. He is a member of *The Language Teacher* Editorial Review Board. He can be contacted at long@dhs.kyutech.ac.jp. Examples of AL language tasks can be found on the author's website (<http://www.genderfluency.com>).



NEAR

北東アジア言語教育学会
Language Education Conference

NEAR Conference
June 1, 2019, Niigata

Join us in Niigata for the 11th annual NEAR conference, the premiere language education event on the Japan Sea coast. The call for papers is open now! <http://nearconference.weebly.com/>

As we enter our 2nd decade, we are proud to announce our new peer-reviewed conference proceedings.

Come to Niigata! Present your work! Publish your results!

2019 MEES

2019 Michinoku English Education Summit

co-sponsored by the Iwate Chapter of JALT & Hachinohe Gakuin University



Sunday, June 23, 2019

At Hachinohe Gakuin University, Hachinohe, Aomori

ELT for What Purpose? Connecting English Learning to Social Contexts

Teach In Tokyo

Searching for overseas teaching opportunities? Where will you be in 2019 and 2020? Why not work in Tokyo which is hosting the Rugby World Cup in 2019 and the Olympic Games in 2020. **International Education Services (IES)**, a language consultancy celebrating 50 years of experience teaching Japanese executives in Japan, is seeking engineering, computer, and business graduates to work as English language instructors. You will teach business English to employees working for our corporate clients and officials from governmental ministries. You will also conduct English lessons for younger learners at Japanese public primary and secondary schools.

Renewable, annual contracts include work visa sponsorship, paid vacations, medical insurance, with a salary around \$30,000. Transportation to and from classes paid by IES. Instructor Openings available throughout 2019.

Please email your cover letter and CV to [<jobs@iesnet.co.jp>](mailto:jobs@iesnet.co.jp)
Additional employment information will be emailed to you.

