

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 readability scores of the transcripts of the listening samples correlated with listening comprehension, were also explored. Participants scored only slightly higher on SL multiple choice tests than on AL tests, despite the SL material containing higher-level vocabulary. There was no significant difference in the performance between low and intermediate proficiency students in this study, and no significant correlation between lexical density and the number of correct answers was found. Notably, higher readability indices of listening sample transcripts did not result in higher comprehension scores, indicating that other variables, such as background knowledge, were possibly more influential. The researcher argues that authentic listening samples that are only slightly more challenging than scripted ones can be found and thereby should be considered for classroom use.

この調査は、日本のEFL学習者の理解度が、スクリプト言語イージーリスニング 基づいたオーテンチェック ランゲージ (AL) ディスカッションとスクリプト ランゲージ (SL) に基づいたクイズで大きく異なるかどうかに関係していました。さらに、上級者と中級者の理解力の違いと、聞き取り課題の転写物のリーダビリティスコアが聞き取り能力とどのように相関しているかを調べました。結果はALとSLのタスクの間には小さな違いもなく、後者はALより平均して間違いが1少ないだけであり、中級者と上級者の間には有意な差はなかった。また、転写物のリーダビリティスコアが高いほど、理解率が低くなることはありませんでした。結果は、オーテンチェック ランゲージの演習は教室でも行われるべきであり、聞き取りの理解度は、ディスカッションの意味を理解しているかを示す指標というよりは、正しい答えを推測して選択 することができるという学生の能力を反映しているかもしれないことを示している。

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 higher or lower readability and lexical textual 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 comprehen-

sion 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. How 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 the readability (the Flesch-Kincaid Grade Level test 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. For this reason and to determine lexical textural density, the texts of each listening

script were analyzed with the Flesch-Kincaid Grade Level test. It was assumed that higher reading scores would reflect more advanced vocabulary usage, which would result in more difficult listening passages. This was a means for comparing SL and AL samples lexically.

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 using the GoldFish and VOA listening material, and only data from those quizzes were analyzed for this study.

Results

An independent *t*-test was used to compare the low and intermediate proficiency learners' listening test results, and the Pearson correlation coefficient was used to measure correlations through the software XLSTAT (an Excel add-on) and StatCalc (AcaStat, 1999). Textual density for the listening samples was analyzed with Textalyser (2014). The results are shown in Tables 1 and 2.

Regarding the first research question concerning listening comprehension scores between AL and SL listening tasks, there was a statistically significant difference between the number of comprehension questions answered correctly, $t(17) = 2.11, p = .04$. However, on average, learners answered only one more question correctly per test on average on SL quizzes than AL quizzes.

The second research question asked whether there were any significant differences in comprehension scores between the low and intermediate proficiency EFL learners. A *t*-test comparing the AL and SL data showed no significant difference, $t(38) = 2.02, p = .98$ (see Tables 3 and 4). On average, learners in both groups scored between nine and 10 points correct on the AL quizzes, and between 10 and 11 points on the SL quizzes.

Regarding the third research question, a Pearson product-moment correlation coefficient was computed to assess the relationship between lexical density and correct answers, but no significant correlation was found, $p = .83, r = -.052$. A Pearson product-moment correlation coefficient was computed to assess the relationship between Flesch Reading Grade levels and correct answers, and a significant correlation was found, $p = .05$, with a small effect size of $r = -.037$. Tables 3 and 4 present scores sorted from low to high, along with the Flesch-Kincaid Grade levels.

Table 1. Descriptive Data from Authentic Language Samples (GoldFish)

Abbreviated Titles of Recordings	Flesch-Kincaid Grade Level	Lexical Density	Words Per Sentence
Buying a House	3.5	52.9%	11.7
Creativity	4.7	57.6%	14.0
Any Questions	3.4	55.5%	15.7
Think/Grow Rich	4.5	49.5%	13.3
Learning/Changing	7.6	48.3%	20.6
Malala 1 & 2	3.6	49.4%	18.3
Same-Sex 1, 2, & 3	1.4	38.3%	15.1
Cigarettes 1,2, & 3	10.1	43.3%	11.2
Climate Change	12.0	55.7%	20.9
Skin Deep	5.9	43.7%	8.9
Doing a Ph.D.	2.8	45.8%	11.1
Average	5.4	48.7%	14.9

Note. Flesch-Kincaid Grade Levels were based on 100 words.

Table 2. Descriptive Data from Scripted Language Samples (VOA)

Abbreviated Titles of Recordings	Flesch-Kincaid Grade Level	Lexical Density	Words Per Sentence
President Trump	10.8	66.0%	17.5
Branding	11.6	67.5%	16.5
Schools Under Pressure	12.0	58.7%	14.9
No Grades	7.0	54.5%	13.8
Scientists Strike Gold	10.9	65.1%	14.1
Let's Talk About Friends	8.8	50.5%	12.2
Afghanistan	9.6	60.1%	16.1
Jane Goodall	5.8	72.9%	17.6
Food Festival	12.0	68.6%	21.8
Average	9.8	64.0%	16.0

Note. Flesch-Kincaid Grade Levels were based on 100 words.

Table 3. Correct Answers and Flesch-Kincaid Grade Level (100 Words) for Intermediate Students

Abbreviated Titles of Recordings	Type/#	Correct Answers	Flesch-Kincaid Grade Level/100
Skin Deep	AL 10	6.7	5.9
Doing a Ph.D.	AL 11	8.2	2.8
Same Sex 1, 2, & 3	AL 7	8.3	1.4
Creativity	AL 2	8.4	4.7
Jane Goodall	SL 8	8.4	5.8
Climate Change	AL 9	8.6	12.0
Think/Grow Rich	AL 4	8.7	4.5
Malala 1 & 2	AL 6	9.0	3.6
Learning/Changing	AL 5	9.1	7.6
Cigarettes 1, 2, & 3	AL 8	9.8	10.1
Let's Talk About Friends	SL 6	9.9	8.8

Abbreviated Titles of Recordings	Type/#	Correct Answers	Flesch-Kincaid Grade Level/100
Buying a House	AL 1	10.2	3.5
Afghanistan	SL 7	10.2	9.6
Schools Under Pressure	SL 3	10.5	12.0
Trump	SL 1	10.6	10.8
Any Questions	AL 3	11.0	3.4
Branding	SL 2	11.5	11.6
Scientists Strike Gold	SL 5	11.8	10.9
No Grades	SL 4	12.3	7.0
Food Festival	SL 3	13.1	12.0

Table 4. Correct Answers and Flesch-Kincaid Grade Level (100 Words) for Beginner Students

Abbreviated Titles of Recordings	Type/#	Correct Answers	Flesch-Kincaid Grade Level/100
Same Sex 1, 2, & 3	AL 7	8.0	1.4
Learning/Changing	AL 5	8.2	7.6
Scientists Strike Gold	SL 5	8.6	10.9
Climate Change	AL 9	8.6	12.0
Cigarettes 1, 2, & 3	AL 8	8.6	10.1
Let's Talk About Friends	SL 6	9.0	8.8
Creativity	AL 2	9.0	4.7
Think/Grow Rich	AL 4	9.1	4.5
Doing a Ph.D.	AL 11	9.2	2.8
Jane Goodall	SL 8	9.2	5.8
Malala 1 & 2	AL 6	9.6	3.6
Any Questions	AL 3	10.0	3.4
Afghanistan	SL 7	10.0	9.6
Skin Deep	AL 10	10.1	5.9

Abbreviated Titles of Recordings	Type/#	Correct Answers	Flesch-Kincaid Grade Level/100
Buying a House	AL 1	10.7	3.5
Branding of Politicians	SL 2	10.7	11.6
No Grades	SL 4	11.3	12.0
Trump	SL 1	11.5	10.8
No Grades	SL 4	11.8	7.0
Food Festival	SL 9	13.0	12.0

Discussion

The results show that there was a consistent, but small, difference in comprehension scores between quizzes for scripted language recordings and authentic language recordings in this study. Comprehension scores between the low-intermediate proficiency EFL learners and the intermediate learners indicated that both levels of students did not find authentic listening more difficult to comprehend than scripted listening in this study.

The reading grade levels of the scripts of the recordings correlated with listening comprehension scores, but only slightly. This could suggest that transcripts that are more difficult to read do not necessarily affect listening comprehension substantially, and that other variables might be responsible. Although the data show that participants scored marginally 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, possibly the lower AL test scores were a result of other features of authentic speech, such as faster speaking rates, more common use of slang, fragmented utterances, repetitions, and rephrasing. 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, as this data suggest, authentic listening materials can be found which are similar in lexical difficulty to scripted materials, then AL tasks should be preferred because they help students both with test results and to develop real-life communication skills.

Another feature which 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.

These results indicate that authentic language recordings exist that even low proficiency learners can comprehend, but that recordings should ideally relate to topics which the listeners are already familiar with. Authentic conversations can be adapted by teachers for classes (Liu, 2016). Student scaffolding can be supported through the provision of videos and scripts, and through communicative activities such as role-playing and script continuation. Examples of such material can be found on the author's website (www.genderfluency.com).

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). Another issue was that the factor of learners' background knowledge was not controlled for.

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 and the *TLT* reviewers and editors for their detailed feedback. I would also like to thank José Cruz for his comments, for his technical help in website design for my website (www.genderfluency.com), and for developing <http://www.genderfluency.com>

www.Goldfish365.com, which provided material for this study.

References

- BBC (2015). *News Report*. 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 Masters degree in English Education and a Specialist Degree in Education Multilingual/Multicultural. He has worked at Kyushu Institute of Technology for 22 years, where he is currently an associate 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.

