Encouraging L2 Science Majors to Talk Using L1 Readers

Sae Matsuda*, Makoto Imura**, and Noriko Nakanishi***

***Setsunan University **Osaka Institute of Technology
***Kobe Gakuin University**

This pilot study attempted to examine whether repetitive input and output with L1 readers can help university science majors improve their speaking skills. The experimental group (14 science majors) had ten 90-minute classes where they read and listened to *Oxford Reading Tree*, followed by a series of practice activities including shadowing, repeating, reading aloud, and a book talk/retelling the story. Three types of online tests—*Progress*, *Versant*, and *OPIc*—were also conducted before and after the treatment, and the results of the pre- and post-tests were compared within the experimental group. The control group (4 science majors) also took the same tests without any treatment. Despite the limited time period, most students in the experimental group showed improvements in some elements measured by the tests mentioned above while the control group failed to maintain their original level.

The Ministry of Education, Culture, Sports, Science and Technology (2015) reported that 90% of senior students in Japanese high school fell into Common European Framework of Reference (CEFR) A1 category or below in terms of speaking and writing. Those beginner-level students enter university and take English again. Students majoring in math and science in particular tend to shy away from English. However, these students are the ones who may have to make full use of English in the future. To make matters worse, English courses provided by universities in Japan are not sufficient to meet their needs and students are not getting enough input or output practice. They often end up simply translating English sentences into Japanese in their limited number of English classes (one or two 90 minute classes per week). Meanwhile, Nishizawa and Yoshioka (2013; 2016) established an extensive reading program at National Institute of Technology, Toyota College using easy materials such as *Oxford Reading Tree* and *Penguin Young Readers* (now known as *Pearson Young Readers*), and after a five-year extensive reading program, their students demonstrated remarkable gains in the TOEIC® test scores (Nishizawa, Yoshioka, & Ito: 2010). Yokoyama (2011) claims that easy levels of *Oxford Reading Tree* and their audio CDs can be an effective tool for speaking practice.

Literature Review

Research has shown that there are various benefits derived from extensive reading (cf. Krashen, 2004; Waring, 2006; Nation, 2013). Day and Bamford (1998) also provided useful summaries of the evidence. Moreover, there is a comprehensive

bibliography available at The Extensive Reading Foundation site (http://erfoundation.org/wordpress/er-bibliography/). There is an abundance of studies showing that extensive reading not only enhances reading competency (cf. Elley, 1991: Mason & Krashen, 1997; Day & Bamford, 1998; Rodrigo, Krashen & Gibbons, 2004) but also improves other skill areas such as writing (cf. Hafiz & Tudor, 1989: Lai, 1993; Mason & Krashen, 1997), vocabulary (cf. Pitts, White, & Krashen, 1989; Day, Omura & Hiramitsu, 1991; Cho & Krashen 1994; Nation, 2001; Horst, 2005), speaking and listening (cf. Hafiz & Tudor, 1989; Cho & Krashen, 1994). Moreover, extensive reading enhances learners' motivation (cf. Elley, 1991; Cho & Krashen, 1994; Takase, 2008). Reading books that learners can easily understand and enjoy is motivating, and Nuttall's (1996) virtuous cycle shows that if they can understand better, they enjoy reading, then read faster and more, and eventually they become fluent readers.

Leveled Readers

According to Nation (2013:37), fluency activities for listening and speaking should have four important features: 1) easy materials, 2) pressure to go faster, 3) message focused, and 4) quantity of practice. Among them, choosing easy materials is the most important. Leveled readers—easy picture books aimed for L1 children—can be suitable materials. They are written with simple vocabulary and syntax. As Takase (2008) and Nishizawa & Yoshioka (2016) report, leveled readers such as *Oxford Reading Tree (ORT)* and *Longman Literacy Land* are appealing to Japanese university EFL students and have been considered as appropriate materials for especially inexperienced readers. More importantly, leveled readers contain another crucial element

for fluency: repetition. Nation (2013) and Kadota (2015) emphasize that repetition is a key to success. *ORT* has plenty of examples of "varied repetition where the same language features re-occur in changing contexts" (Nation, 2013: 38). Then can repetitive input and output using *ORT* help improve science majors' speaking skills?

CEFR-J and Speaking Descriptors

Finding that "the population of Japanese EFL learners skews towards the lower levels," Negishi (2012:105) subdivided CEFR A1-B2 levels and created the modified Japanese version (CEFR-J). The following are excerpts from the CEFR-J speaking "can-do" descriptors:

<Presentation>

A2.1   I can give a brief talk about familiar topics (e.g. my school and my neighborhood) supported by visual aids such as photos, pictures, and maps, using a series of simple phrases and sentences.

A2.2   I can make a short speech on topics directly related to my everyday life (e.g. myself, my school, my neighborhood) with the use of visual aids such as photos, pictures, and maps, using a series of simple words and phrases and sentences.

B1.2   I can give an outline or list the main points of a short story or a short newspaper article with some fluency, adding my own feelings and ideas.

<Interaction>

A2.1   I can get across basic information and exchange simple opinions, using pictures or objects to help me.

The presentation descriptors above include using pictures, simple phrases and sentences, and a short story. As for the interaction skills, students are expected to get across basic information and exchange simple opinions, assisted by pictures. Wouldn't it be possible to show science majors a way to improve their English speaking ability using L1 picture books and get them to speak English that way?

Extensive Reading and Speaking

Measuring learners' speaking ability remains a big challenge. Hato, Takei, Healy, Kamizawa, and Ito (2015) tried to assess examinees' speaking ability by conducting an original interview test. However, they concluded that it took too much time and energy, and it would be impractical as they designed it to carry on the same procedures with limited human resources in a national university. Now the government and universities are looking into outside sources for testing speaking ability. *Versant*—an online speaking test—has been used to electronically measure learners' speaking ability before and after studying abroad (cf. Yoshida, Kagata & Ikuma, 2012; Shimizu, Kirimura & Nozawa, 2014). Yet, to the authors' knowledge, no research has been conducted using online speaking tests to measure effects of extensive reading on spoken output. Although Nakanishi & Ueda (2011) reported positive effects of reading aloud and shadowing on *Secondary Level English Proficiency Test* (*SLEP*) using extensive reading books, the participants' speaking skills were not measured.

Takeuchi (2003) claims that reading aloud is an effective way to learn a language. Those science majors who are lacking output practice could benefit from reading aloud. It is argued that the language is internalized when learners read aloud after understanding the content and pronunciation.

Kadota (2015: 64-66) points out two effects of reading aloud—automatization of written lexical access and internalization of new items—and two effects of shadowing—automatization of speech perception and internalization of new items.

The Pilot Study

The Purpose of the Study

The pilot study was conducted to explore whether repetitive input and output with L1 readers enhance L2 science majors' speaking skills, and if so, whether the progress can be measured by online tests.

Participants

The pilot study project was advertised through the network of the Faculty of Science and Engineering. Two information meetings were held in June, which drew twenty-three students. Some students attended voluntarily while others were prompted by their seminar teachers. At the information meetings, the students learned the purpose of the study and the content of the activities. After hearing the explanation, the students submitted their available time slots, and the fifth period on Monday was selected because it was the only slot when most students were available. Due to time conflicts, four students had to move to the control group while five students decided to withdraw from the project. Finally, in the beginning of September, 2016, the project launched with fifteen students in the experimental group and five students in the control group. Table 1 shows the number of participants, their majors, and the year of study.

All the first-year students (indicated with two asterisks) were taking two 90-minute English classes a week as part of their core curriculum while two second-year students

Table 1. Number of Participants

	Department	1st year	2nd year	3rd year
Experimental Group	Civil & Environmental Engineering		1	
	Electrical & Electronic Engineering	6**	1*	
	Life Science	2**	2	2
	Mechanical Engineering		1*	
Control Group	Architecture	3**		
	Electrical & Electronic Engineering	2**		

(indicated with one asterisk) were taking one weekly 90-minute English class. The rest of the students (without asterisks) were not taking any English classes during the treatment period. One student from the control group dropped out after taking the *Progress* pre-test while one student from the experimental group withdrew in Week 5 because she was tied up with her studies; therefore, their data were removed from the results.

Method

Instruments

The following excerpts are the official descriptions of the online tests used in this study:

1) Progress (Pearson)

Progress is an online, integrated skills English language proficiency test package reporting scores on the Global Scale of English for listening, reading, writing, speaking, grammar and vocabulary. The tests feature a combination of both adaptive and linear sections which include multiple versions of the linear tests and a large number of questions in the item bank for the adaptive section.

(http://media.pearsoncmg.com/intl/elt/ioki/gse-global-scale-of-english/course_docs/about-the-test.pdf)

2) OPIc (Global 8)

The *OPIc* is an internet-delivered test which provides valid and reliable oral proficiency testing on a large scale. The computer-delivered assessment emulates the "live" OPI, but the delivery of questions is through a carefully designed computer program and via a virtual avatar, allowing the test to be taken on demand and at a time convenient to the candidate and proctor.

(https://www.languagetesting.com/oral-proficiency-interview-by-computer-opic)

3) Versant (Pearson)

The *Versant* testing system, based on the patented Ordinate® technology, uses a speech processing system that is specifically designed to analyze speech from native and non-native speakers of the language tested. In addition to recognizing words, the system also locates and evaluates relevant segments, syllables, and phrases in speech. The *Versant* testing system then uses statistical modeling techniques to assess the spoken performance.

(https://www.versanttests.com/technologyresearch)

Surveys & Interviews

Three surveys (at the beginning, in the middle, and at the end of the term) and two interviews (midterm and end of the term) were also conducted, however, the details cannot be included in this paper due to space limitations.

Procedures

After students took three kinds of online pre-tests (*Progress*, *OPIc*, and *Versant*) in Weeks 1-3, the experimental group had ten 90-minute classes where they engaged in the previously described reading and speaking activities. In Week 3 after the *Versant* test, an orientation was provided in which the instructors (authors) explained the procedures and demonstrated how to do shadowing and repeating activities. In Weeks 14 and 15 they took online post-tests. In Week 16 they received their results and feedback. The main materials were *Oxford Reading Tree*, starting with Stage 1+ set. With OUP permission, the audio files had been installed into the CALL system prior to the training. For the treatment, the students read the books while listening to the audio, then they practiced shadowing and repeating, and after that, they read the book aloud to themselves. After they read three books, they recorded a book talk, or later retold the story about one of the three books. One of the instructors administered the training sessions alone from Week 4 to the end, but there were no explicit instructions provided about pronunciation,

grammar, etc. When the students finished reading all the books on the level they were working on, they came to the instructor and performed a book talk or retold the story to her using their favorite book from that level. Mid-term, one-to-one conferences were also conducted in Week 8 to see how the students were doing. On the other hand, the control group only took pre- and post-tests without receiving any treatment.

Results

The Amount of Input and Output

As shown in Table 2, in ten weeks, the students read 68 books (4,244 words) on average. When the repetition was counted, they read/spoke 272 books (16,976 words) on average. In addition, they gave 6.7 book talks and engaged in retelling 6.9 times on average. By the end of the term, the slowest achiever –who missed 6 classes– had read 37 books (1,825 words) and finished in the middle of Stage 2 while the most diligent student read 96 books (6,002 words) and finished at Stage 4.

Assessment

1) Progress

Figure 1 below displays the average scores of *Progress* pre- and post-tests in the experimental group (*n* = 14) and in the control group (*n* = 4). The bottom score "<10" was calculated as "9" in this study to distinguish it from "10." In order to examine the

Table 2. The Amount of Input and Output

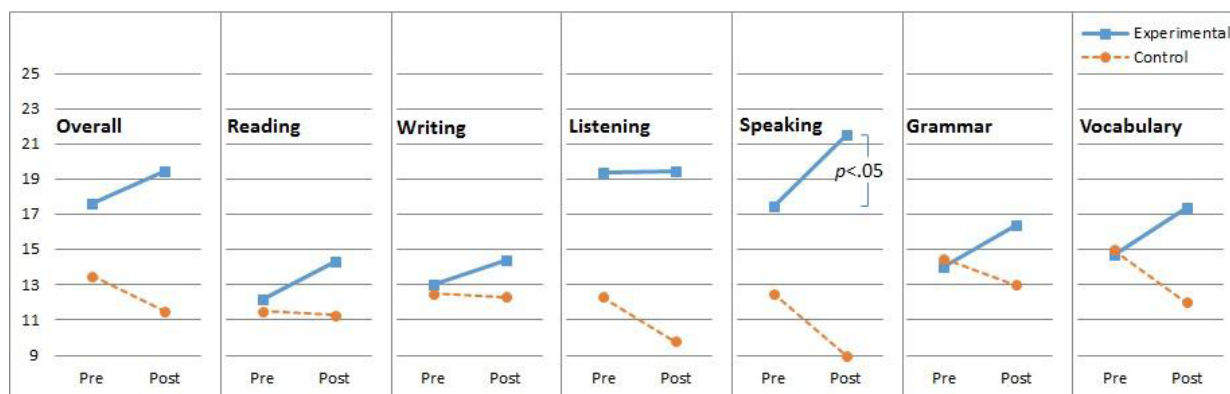|	Type (Repetition Uncounted)		Token (Repetition Counted)	
	Books	Words	Books	Words
Mean	68	4,244	272	16,976
Max	96	6,002	384	24,008
Min	37	1,825	148	7,300

Figure 1. The average scores of Progress pre- and post- tests.

experimental group's progress, a paired t-test was conducted. A statistically significant result was demonstrated in speaking (t (13) = 2.28, $p < .05$).

2) *OPIc*

One student from the experimental group missed the post test, thus her data were excluded from the analysis. Figure 2 shows the transition of levels demonstrated by the experimental group ($n = 13$) and the control group ($n = 4$).

The post-test results revealed that six students out of 13 in the experimental group reached one level higher than their original level. However, all the five students who started with NH or IL stayed in the

same level. Meanwhile, one student in the control group moved from NM to NH, but all the other three students maintained the same level.

3) *Versant*

Two students from each group received "not scored" results: i.e. "The candidate's responses may have been in a language other than English, produced with poor/ unintelligible pronunciation, or irrelevant," according to Pearson. Thus, a total of four students were removed. Furthermore, one student in the experimental group failed to take the post-test. Therefore, the results of 13 students in total (11 from the experimental group and 2 from the control group) were analyzed. Figure 3 indicates

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	C1	C2	C3	C4
IL												Post Pre	Post Pre				
NH					post	post	post	post	Post Pre	Post Pre	Post Pre					post	Post Pre
NM	post	post	Post Pre	Post Pre	Pre	Pre	Pre	Pre							Post Pre	Pre	
NL	Pre	Pre												Post Pre			

Figure 2. The transition of OPIc levels from pre- to post- test. The upper part of the figure indicates higher levels; i.e. IL=intermediate-low, NH=novice high, NM=novice-mid, and NL=novice-low.
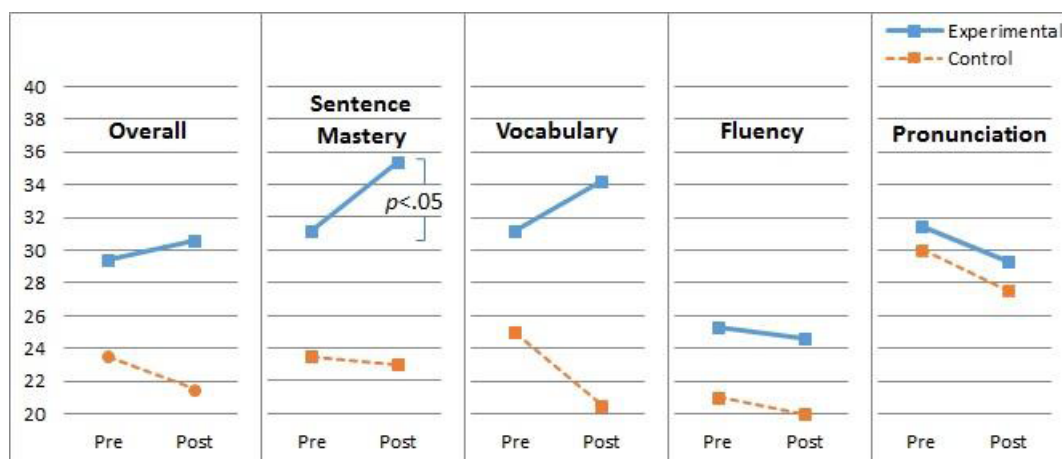
Figure 3. The average scores of Versant pre- and post-tests.

the average scores of *Versant* pre- and post-tests. For the experimental group, a paired t test was conducted.

Versant yielded both positive and negative results. In the post-test, the experimental group scored higher in sentence mastery (t (10) = 3.12, $p < .05$) than in the pre-test, but any statistically significant result was not found in other scores. One of the two students in the control group got one point higher on vocabulary, but otherwise the group didn't show any improvement.

Discussion

The results of the pre- and post-tests within the experimental group yielded a statistically significant positive change in speaking (*Progress*) and sentence mastery (*Versant*). At the beginning of the treatment, it seemed very challenging for the lower-level students to talk about the books they read, and many students took a long time constructing English sentences before they recorded their stories. When demonstrating a book talk or retelling the story to the instructor, they rehearsed many times before they came. Nevertheless, their careful attention to sentence making may have worked positively on sentence mastery measured by *Versant*.

The *Versant* results were a bit surprising since the authors expected that if the treatment is effective, it might work positively for fluency or pronunciation first. A possible reason for lower scores on pronunciation and fluency may be due to lack of attentive listening or speaking. It can also be an adverse effect of the book talk and the story retelling mentioned above. In other words, the students may have become too careful about their speech and spoke too slowly, which could have hampered fluency building. During the treatment, no feedback was given about pronunciation; thus the students had no idea if their speech was appropriate linguistically or contextually.

OPIc results implied that the treatment this time worked best for the beginner-level learners. Among the different levels of students, the NL and NM level students were more likely to demonstrate progress than the NH or IL level students. According to Tschirner (2012), novice (low, mid, high), intermediate-low, and intermediate-mid levels on *OPIc* are equivalent to CEFR A1, A2, and B1 respectively. In other words, NL, NM, and NH levels are all in CEFR A1 category. Naturally, it is easier for novice students to move up within A1 category. In contrast, IL students (CEFR A2) cannot

advance to IM (CEFR B1) that easily. The simple treatment in ten weeks was just not sufficient.

Anecdotally, it was observed that those students who participated actively enjoyed reading aloud (even imitated dog sounds happily and loudly) and made remarkable progress on many elements of the three tests. On the other hand, the students who spoke quietly and didn't change prosody or intonation very much failed to receive positive results. Nevertheless, everyone made a few improvements on a test or two at least, and they became less reluctant to talk after the treatment. It was found that most of the students, who were used to the grammar-translation type of learning, enjoyed the reading, listening, and speaking exercises using leveled readers.

In turn, some limitations of the study should be taken into account. In order to recruit students, the authors explained research objectives to prospective participants, however, it might have directed their attention to speaking and possibly affected the results. The participation was based on a volunteer basis, and the small size of the project made it difficult to form a proper control group. There were students taking one or two English classes a week, which might have influenced their performance on post-tests. Although there was no other choice, meeting in the 5th period (16:40- 18:10) on Mondays may not have been preferable, and it is possible that the fatigue factor affected their outcome. The repetition of tasks was at times daunting, and some students lost concentration in the middle of the class.

For further studies, the elements that constitute speaking skills should be scrutinized. What factors determine speaking skills?

Would it be the ability to keep talking, or to organize thoughts? What about the rate of speech, prosodic features, word variety, sentence structures? What about the use of gap fillers and other circumvention strategies when they get stuck? Should their individual differences such as anxiety, willingness to communicate (WTC), and motivation be taken into consideration?

The next step will be narrowing down the elements to measure and reconsider the tasks to assign. In particular, more effective ways to improve fluency and pronunciation should be taken into account. It is also necessary to look into individual performance in more detail rather than just depending on the test results. By analyzing the students' recorded speeches, lexico-grammatical features as well as phonetic features could be observed.

Conclusion

This pilot study attempted to explore whether the repetitive input and output bring measurable improvements in students' speaking skills, and whether the results could be measured by three online tests. The speaking skills measured by *Progress* and sentence mastery assessed by *Versant* yielded statistically significant improvements in the experimental group. As mentioned at the beginning of this paper, most students, particularly science students, enter tertiary education at CEFR A1 level, and they are the most unwilling to speak English. What they should ideally aim at achieving is B2 level because it is the level where they can understand the main ideas of complex texts in both concrete and abstract topics, including technical discussions in their field of specialization. It seems a long way to go, but practicing with simple ER books such as *Oxford Reading Tree* may provide a new approach to studying

English and give them a nudge to try. That will be the first step.

References

Cho, K. S., & Krashen, S. (1994). Acquisition of vocabulary from the Sweet Valley Kids Series: Adult ESL acquisition. *Journal of Reading, 37*(8). 662-667.

Day, R., & Bamford, J. (1998). *Extensive Reading in the second language classroom*. Cambridge: Cambridge University Press.

Day, R., Omura, C., & Hiramitsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7*(2), 541-551.

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*(3), 375-411.

The Extensive Reading Foundation (ERF). *ER Bibliography*. Retrieved from http://www.erfoundation.org/erf/

Hafiz, F. M. & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal, 43*(1), 4-13.

Hato, Y., Takei, T., Healy, S., Kamizawa, K., & Ito, K. (2015). *Introducing a speaking test of English into the university entrance examination: Evaluation of a computer–based pilot test*. Paper presented at the 54th JACET national convention.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review, 61*, 355-382.

Kadota, S. (2015). *Shadowing ondoku to Eigo communication no kagaku [Science of communication in English]*. Tokyo: Cosmo Pier.

Krashen, S. (2004). *The power of reading—Insights from the research*. Westport, CT: Libraries Unlimited Inc.

Lai, F. (1993). The effects of summer reading course on reading and writing skills. *System, 21*(1), 87-100.

Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System, 25*(1), 91-102.

Ministry of Education, Culture, Sports, Science and Technology (2015). *Heisei 26 nendo eigo kyouikukaizen notameno eigoryoku chousajigyou houkoku [The report on the investigation for improving English skills 2014]*. Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikoku-go/1358258.htm

Nakanishi, T., & Ueda, A. (2011). Extensive reading and the effect of shadowing. *Reading in a Foreign Language, 23*(1), 1-16.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. P. (2013). *What should every EFL teacher know?* Seoul, South Korea: Compass Publishing.

Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2nd ed.). Oxford: Heinemann.

Negishi, M. (2012). *The Development of the CEFR-J: Where we are, Where we are going*. Retrieved from http://www.tufs.ac.jp/common/fs/ilr/EU_kaken/_userdata/negishi2.pdf

Nishizawa, H., Yoshioka, T. & Ito, K. (2010). The effectiveness of a 5-year extensive reading program. *JERA Bulletin, 4*, 1-14.

Nishizawa, H., & Yoshioka, T. (2013). ER as virtual ESL environment for EFL learners. *The Second Extensive Reading World Congress Proceedings*, 126-137.

Nishizawa, H., & Yoshioka, T. (2016). Longitudinal case study of a 7-year long ER program. *The Third Extensive Reading World Congress Proceedings*, 28-40.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the clockwork orange study using second language acquirers. *Reading in a Foreign Language, 5*, 271-275.

Rodrigo, V., Krashen, S., & Gribbons, B. (2004). The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. *System, 32*(1), 53-60.

Shimizu, Y., Kirimura, R., & Nozawa, T. (2014). Effects of short-term study abroad programs for Economics students: Based on analysis of speaking tests. *Ritsumeikan Koto Kyoiku Kenkyu, 14*, 91-102.

Takase, A. (2008). The two most critical tips for a successful extensive reading program. *Kinki University English Journal, No. 1*, 119-136.

Takeuchi, O. (2003). *Yoriyoi gaikoku-gogakushuho wo motomete [Searching for better language learning strategies: Studies on good language learners in the Japanese FL context]*. Tokyo: Shohakusha.

Tschirner, E. (2012). *Assessing evidence of validity of assigning CEFR ratings to the ACTFL Oral Proficiency Interview (OPI) and the Oral Proficiency Interview by computer (OPIc)*. Retrieved from http://global8.or.jp/OPIc%20CEFR%20Study%20Final%20Report%20pdf.pdf

Waring, R. (2006). Why Extensive Reading should be an indispensable part of all language programmes. *The Language Teacher, 30*(7), 44-47.

Yokoyama, K. (2011). Reading British literacy textbooks: An introduction to Oxford Reading Tree. *Gunma-Kosen Review, 30*, 101-112.

Yoshida, H., Kagata, T., & Ikuma, Y. (2012). Implementing Versant: Spoken English test on program for promoting university education and student support based on different levels of English proficiency. (in Japanese). *The Bulletin of the English Society, 57*, 35-62.