Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education. New Zealand: Asia-Pacific Society for Computers in Education

# Effectiveness of a Learning Design Combining Summary-speaking Self-study Using Mobile Application with Paired Reflection on Learners' Speaking Process

**Kae NAKAYA[a*] & Masao MUROTA[b]**
[a]*Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Japan*
[b]*Institute for Liberal Arts, Tokyo Institute of Technology, Japan*
*knakaya@mr.hum.titech.ac.jp

**Abstract:** The authors earlier developed the Mobile Applications for Scaffolded autonomous summary speaking Task (MAST) for supporting learners' self-study of English speaking. Although MAST enables learners to conduct summary-speaking tasks without teachers' support and improves their oral fluency, it cannot offer enough feedback to enable learners to test whether the listener can comprehend their oral output, which is also necessary for learners to acquire second language. Therefore, we proposed a learning design that combines self-study using MAST and reflection activities in pair work in an English class. This research's objective was to investigate the effectiveness of the learning design for learners' oral performance and to clarify the role of self-study and reflection activities for development of learners' speaking skills. After analysis, we concluded that the learning design was effective for improving learners' oral performances. We also discussed the relation between the oral performances and learners' cognitive process in our learning design.

**Keywords:** Computer assisted language learning, oral task, summary speaking, pair work, reflection, oral performance

## 1. Introduction

Because globalization continues to accelerate, people whose native language is not English face the important issue of improving their English skills. However, self-study especially practice English-speaking is difficult for English as a Foreign Language (EFL) learners without a teacher's support.

To address this problem, the present authors developed Mobile Applications for Scaffolded autonomous summary speaking Task (MAST) (Nakaya and Murota, 2016a). In this system, learners speak an English newspaper article's summary in English through self-study. The feature of the system is to offer scaffolding practice for speaking its summary. By conducting the practice, learners could understand the main points of the article and necessary words of its summary. We found that MAST enables learners to practice speaking English on their own and to improve their oral fluency.

Despite the effectiveness, learners needed more help to reflect on their speech. Using MAST, learners could check their recorded summary-speaking and compare it with a sample summary as a reflection activity. In the activity, the learners tended to reflect on their speech in terms of only fluency. Swain (2005) stated the importance of a peer's feedback for second language (L2) acquisition. In order for learners to test whether a listener can comprehend their oral output and to seek better expression, feedback from other people seemed to be needed.

Therefore, in a previous study, we proposed a learning design that combined self-study using MAST and pair work in an English class (Nakaya and Murota, 2016b). With this design, learners practiced English summary speaking through self-study at home using MAST, and after that, they told the summary to a peer and next, reflected on it with the peer, using a worksheet designed for an English class. The worksheet facilitated the peer in clarifying and discussing improvements of learners' summary-speaking step by step.

Although we found that the learning design above increased learners' motivation for speaking English outside the classroom (Nakaya and Murota, 2016b), it remains unclear that why the combined learning design is effective. Consequently, the objective of this research is to clarify the roles of self-study and subsequent pair-reflection activities by evaluating the learning design's effectiveness for learners' oral performance in detail, targeting Japanese undergraduate and graduate students.

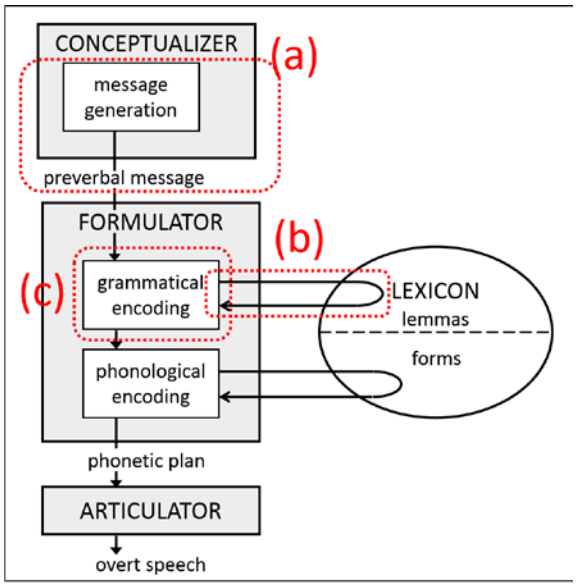


Figure 1. Speech Production Model Based on Levelt (1989).

## 2. Background

### *2.1 MAST's Effectiveness on the Speaking Process*

In designing a method for supporting learners' speaking skills, considering the influence of oral tasks' features on learners' cognitive process is important. Skehan (2009) attempted to explicate the relation among learners' oral performances, features of oral tasks, and learners' speaking processes based on Levelt's model (1989). But before considering a MAST design based on Skehan's suggestion, we first explain Levelt's speech production model.

Levelt's model (Figure 1) illustrates how, when speaking, people process information in three main stages: Generating preverbal messages in Conceptualizer (Figure 1-a), retrieving lexical information from Lexicon (Figure 1-b), and building syntactic structure in Formulator (Figure 1-c)). First, Conceptualizer generates a preverbal message and sends it to Formulator. Second, Formulator retrieves the necessary words from Lexicon. Third, Formulator processes the information grammatically. Fourth, Formulator encodes the structure into a phonetic plan. Fifth, Articulator converts the phonetic plan into audible sound.

According to Skehan (2009), if the information that learners process in Conceptualizer is concrete and easy, they can pay more attention to process in Formulator. In addition, readiness for using necessary lexical items enables learners to retrieve them smoothly from Lexicon. Therefore, when learners' cognitive load for generating a preverbal message and retrieving words decreases, they can allocate cognitive resources to grammatical encoding.

Scaffolding practice in MAST was developed according to Levelt's model and Skehan's discussion. MAST offers scaffolding that enables learners to understand a text's main points and what kinds of words they can use to summarize it aloud. That is, learners do not have to pay much attention to generating a preverbal message (Figure 1-a) and retrieving the necessary words (Figure 1-b); thus they can focus on grammatical encoding (Figure 1-c). By repeating focused grammatical encoding, automaticity of processing linguistic form increases, in turn resulting in improved fluency (Nakaya and Murota, 2016a).

### *2.2 The Role of Reflection in a Learner's Speech with a Peer*

Swain's (2005) output hypothesis indicated peers' importance in L2 learning. If a learner has a partner when speaking English, the learner can test through their peer's feedback whether her or his knowledge construction of L2 is correct. This process provides the learner opportunities to modify her or his output, resulting in L2 acquisition.

If the partner is a native speaker, the learner can obtain immediate feedback through interaction (e.g., recast; the native speaker correctly modifies what a learner has said). However, in classrooms where English is not a native language, most students are non-native speakers, and thus for a peer to provide immediate and appropriate feedback is often difficult.

Therefore, we designed a reflection activity on a learner's summary speech with a peer in a classroom. After a learner speaks a summary in English, the learner and peer, through discussion based on a worksheet, gradually clarify the following points; (1) what the peer could understand during the speech, (2) what the peer could not understand, and (3) improvements to the learner's speech, for example, other words, phrases or additional information to help the peer better understand the summary's content. Following this procedure, the pair can consider together what and how the learner needs to improve in the speech so that the learner obtains valuable feedback in spite of pair work between non-native learners.

### *2.3 Literature Review of CALL Researches*

In terms of supporting learners in improving oral performance, computer-assisted language learning (CALL) researchers evaluated effectiveness of technologies or learning designs that combine self-study using technologies outside the classroom and activities in the classroom.

As for effectiveness of technologies, many researchers clarified how and why technologies contribute to improvements of learners' speaking skills. For example, spectrograms or pitch contours of a learner and a native speaker using automatic speech recognition could work as a visual feedback and improve learners' pronunciation (e.g. Hardison, 2004; Olson, 2014). By offering such visual feedback, learners could recognize the gap between a native speaker's pronunciation and the learner's own pronunciation (Olson, 2014). Regarding synchronous computer-mediated-communication (SCMC), it can create learning environment similar to face-to-face communication. Previous researches have evaluated its effectiveness such as increasing amount of learners' output (e.g. Abrams, 2003) and negotiation for meanings (e.g. Yanguas, 2010). Other researches have explored cognitive mechanism when using SCMC and influence on learners' oral performance (e.g. Payne and Ross, 2005).

Another studies which focused on the learning designs have tended to demonstrate the effectiveness on learners' perception or on their overall performance. Students conducted some learning activities using technologies such as watching English videos, and worked on oral tasks in pairs or groups in class. The combined learning design could ameliorate not only learners' performance (e.g. Hung, 2015) but also their perception (e.g. Cheng et al., 2010).

In contrast, the final goal of our research is to clarify the role of self-study using technologies and the role of classroom activities, and how technologies support and maximize effectiveness of face-to-face learning activities. To achieve the goal, we investigated effectiveness of the combined learning design on learners' oral performance based on Levelt's model (Levelt, 1989).

## 3. Learning Design

In our learning design, learners first do self-study and practice English summary-speaking using MAST; then they conduct pair work in English class, as described below.

### *3.1 Learning Procedure using MAST by Self-study*

Figure 2 shows screens and the learning procedure for MAST. A learner reads an English newspaper article (Figure 2-a) and summarizes it aloud (Figure 2-e). After reading the article and before the summary-speaking task, the learner conducts three scaffolding practices (Figure 2 b-d). In a week, the

learner completes the five tasks (Figure 2 a-e) for one article and can repeat all the tasks as many times as she or he wants.

Details of MAST procedure are as follows: First, a learner reads an English newspaper article. During this activity, the learner can record some words (maximum of five) from the article by tapping on them (a). Words recorded in this step are shown on the screen during summary-speaking tasks (c - e), so the learner can refer to the words for the oral summary. Second, MAST offers Short question and answer practice (b). MAST vocalizes a question related to the article's summary points, and the learner immediately voices the answer. After that, MAST voices a sample answer and then offers the next question. The learner repeats the pseudo-interaction five or six times. Third, the learner conducts two kinds of scaffolded summarizing tasks (c and d). During the learner's summary speaking, MAST screen shows the previously recorded vocabulary list, all the sentences, and the picture in (c), the vocabulary list and the picture in (d). Finally, the learner summarizes aloud referring only to the vocabulary list (e).

For using MAST, we designed two kinds of scaffolding to decrease learners' cognitive load for summarizing aloud. One scaffold is Short question and answer practice (b). By listening to the question, answering it, and listening to the next comment, learners can test their comprehension of the article, clarify its main points, and obtain hints for modifying the summary. Moreover, learners can conduct this practice as a pseudo-interactive conversation with a virtual tutor. The screen shows a female picture, and her face changes according to the dialog. Learners must immediately answer the tutor's question. This procedure possibly decreases learners' language anxiety and increases their motivation. The other kind of scaffolding is Fading summary speaking tasks (c and d). The article's information on the screen fades step by step, so learners can try more difficult summary speaking in small steps, which leads to the main summary-speaking (e).

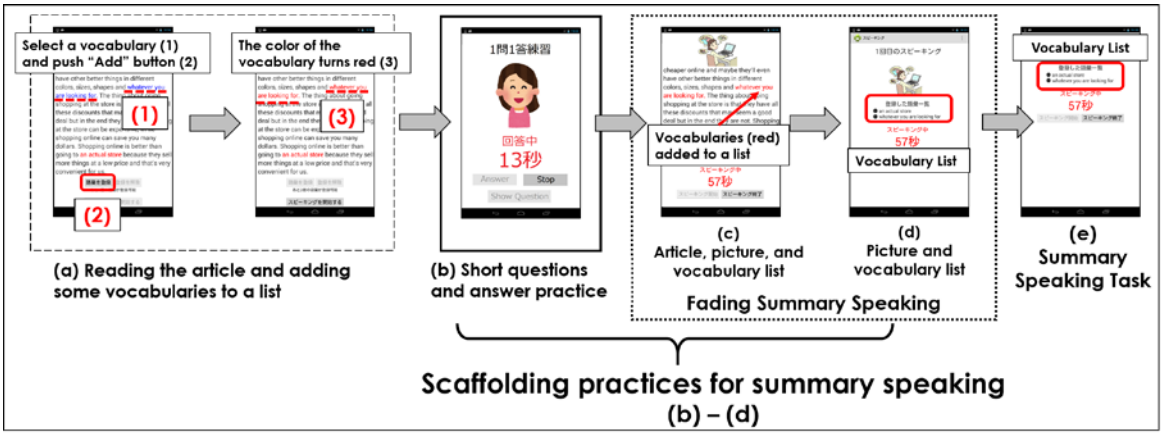


Figure 2. The Screen and Procedure of MAST.

### *3.2 Procedure of Pair Work in an English Class Using a Worksheet*

Figure 3 is the worksheet designed for pair work. The picture (Figure 3-a) and the description (Figure 3 b-e) were written by one of the learners and a peer during an experiment described in section 4.

The detailed procedure using the worksheet is as follows. First, learners conduct a speaking task. A learner speaks a summary, which she or he practiced with MAST at home, to a peer. During the summary, the peer-listener draws a picture to express what the peer understands (a). Second, the pair reflects on the summary. They discuss what the peer-listener could and could not understand based on the picture. Third, the learner writes what the peer-listener could not understand (b) and detailed scripts that she or he spoke actually (c). Fourth, they think of other words or phrases that would help the peer-listener better comprehend the summary, and note the improved scripts (d). If the pair gets other feedback from the other pair or the teacher, they note it in (e).

The worksheet has two significant features. First, the picture drawn by the peer-listener helps the pair visualize the speaker's learning achievement. By observing the picture, the speaker can confirm how much of the summary the listener could understand. This might provide objective feedback to the learner and lead to discussion on effective improvement of the summary. Second, the worksheet facilitates the pair's reflection on the summary speaking and thinking step by step of detailed improvements.

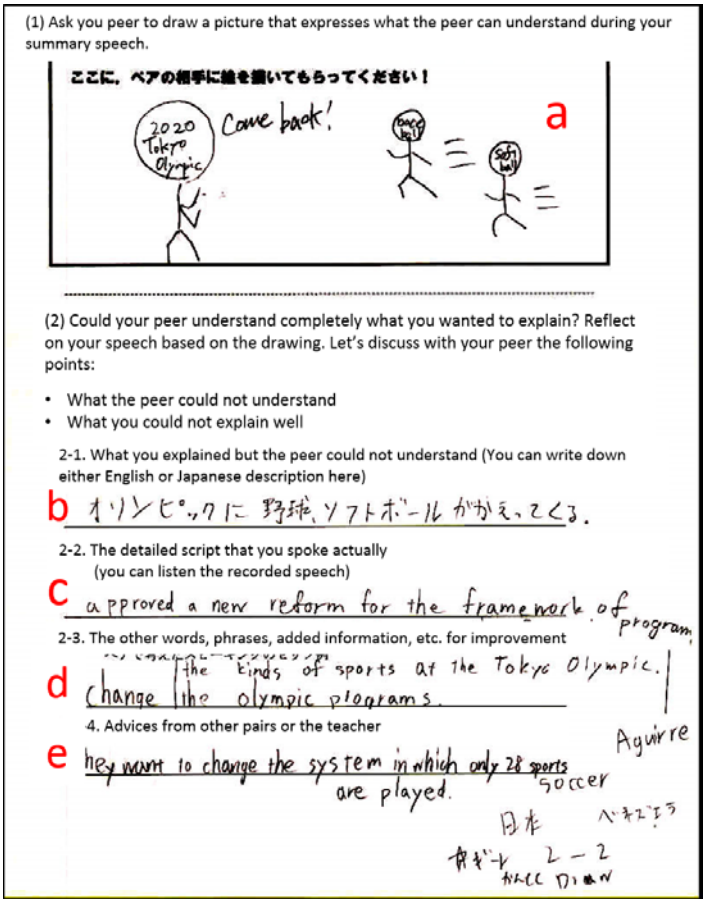


Figure 3. Worksheet for Reflection in Pair Work (instruction translated into English).

## 4. Outline of the Experiment

### *4.1 Objective*

This experiment's objective was to investigate the learning design's effectiveness on learners' oral performance. We implemented pre- and post-tests with speaking tasks, transcribed the speaking tests, and evaluated them, using measures for oral performance. We describe the measures in section 4.3 below.

To clarify the effectiveness of a combination of self-study using MAST and the reflection activity during classroom pair work, we compared results of the learning design in this research with results of the previous learning design that offered neither scaffolding in self-study, nor the reflection activity with a peer (Nakaya and Murota, 2015). In the previous learning design, learners practiced summary speaking without scaffolding practice, using a mobile application at home, and then they explained the summary in English to a peer, but reflected on the summary independently. Therefore, we could utilize the previous learning design's data as a control group. We attempted to clarify its effectiveness by observing not only test scores, but also learners' output during pair work.

### *4.2 Schedule*

We conducted an experiment for four weeks (29th October to 26th November, 2015). On the first day, we explained the experiment to learners and lent everyone a tablet PC (Nexus 7) on which MAST was installed. Using MAST at home for six days, learners practiced English summary speaking; on the seventh day, they conducted pair work in an English class. Learners were 25 Japanese first-year undergraduate students majoring in computer science. Two learners did not participate in the pre-test, and the other two learners had trouble in the pre- or post-test. Therefore, we excluded their data.

In June 2015, the previous learning design's experiment was conducted in a different class of the same university, using the same framework. Six learners did not use the previous MAST at all, and one learner had trouble in the pre-test. Therefore, we excluded their data from analysis. For detailed information, please see Nakaya and Murota (2015).

### *4.3 Data Collection*

As for learners' oral performance, many researchers have measured three items: Complexity (Structural and Lexical), Fluency (Speech rate, Repair fluency, and Silence), and Accuracy (e.g., Tavakoli and Skehan, 2005; Mehnert, 1998). Based on these studies, we transcribed learners' speeches in pre- and post-tests and calculated the following scores.

For fluency, we evaluated (1) Speech rate measured by counting the number of non-repeated words per minute; (2) Repair fluency measured by counting repetitions of exact words, syllables, or phrases, corrections, and partial repeats, and (3) Silence measured by counting silences that last 0.4 seconds or more. Higher Speech rate, and lower Repair fluency and Silence mean improvements of learners' oral fluency.

For complexity, we evaluated (1) Structural complexity measured by counting non-repeated words per the Analysis of Speech Unit (AS-unit) (Foster et. al, 2000), and (2) Lexical complexity measured by counting the ratio of unique words. Higher Structural complexity and Lexical complexity show amelioration in a learner's oral complexity.

For accuracy, we evaluated the ratio of error-free AS-units. A higher score of accuracy means that a learner can speak English with less errors.

When observing learners' reflections, we counted how many noted detailed improvement (other words, phrases, or added information) on the worksheet. Figure 3 displays a typical example describing "other phrases." The learner spoke "…approved a new reform for the framework of program", and the learner and a peer thought up the other phrase: "…change the kinds of sports at the Tokyo Olympic." They tried to use another easier phrase to help the peer's better comprehension of the summary. If a learner wrote more than one kind of improvement, we counted both.

## 5. Results and Discussion

We analyzed all data on learners' oral performances by mixed two-way repeated measures ANOVA using two variables. The between-subjects factor was "learning methods," which had two levels (the current learning design and the previous learning design), and the within-subjects factor was "test term," which had two levels (pre- and post-test). Table 1 shows the results.

### *5.1 Results in Complexity*

Figure 4 displays results for Structural complexity and Lexical complexity.

For Structural complexity, marginal differences were revealed in the "learning method" factor ($F(1,34)=3.288$, $p<0.1$), "test term" factor ($F(1,34)=3.436$, $p<0.1$), and interaction ($F(1,34) = 3.942$, $p < 0.1$). Therefore, we conducted tests of simple main effects. Results for the "test term" factor in the experimental group showed significant differences ($F(1,34) = 7.370$, $p < .05$), and Structural complexity in the post-test was higher than in the pre-test, as shown in Figure 4-a. Results for "learning method" in the pre-test showed significant differences ($F(1,68) = 6.916$, $p < .05$) and the control group's scores were higher than the experimental group's scores.

Lexical complexity showed no significant differences, as shown in Figure 4-b.

### *5.2 Results in Fluency Scores*

For Repair fluency, significant differences were shown on the "learning method" factor ($F(1,34) = 4.430$, $p < .05$). The control group's scores were higher than the experimental group's scores. Figure 5-a shows the results.

For Speech rate, significant differences were shown on the "learning method" factor ($F(1,34) = 4.454$, $p < .05$) and the "test term" factor ($F(1,34) = 25.764$, $p < .001$). These results show that both groups' scores improved, but the experimental group's scores were higher than the control group's scores. Figure 5-b shows the results.

For Silence, significant differences were shown on the "test term" factor ($F(1,34) = 1211.925$, $p<.001$) and marginal differences were shown on interaction ($F(1,34) = 57.407$, $p < .1$). Therefore, we conducted tests of simple main effects. Results for the "test term" factor in the experimental group and the control group showed significant differences ($F(1,34) = 44.871$, $p < .001$ for the experimental group; $F(1,34) = 18.524$, $p < .001$ for the control group), that is, both groups' scores improved. Figure 5-c shows the results

It is worth noting that improved Speech rate and Silence scores of both groups showed that summary-speaking by self-study and telling the summary to a peer, which were conducted by both groups, were effective for these two scores. The improvements were shown on Figure 5-b and c.

In addition, note that statistically significant differences between learning methods in Speech rate and Repair fluency just showed the differences between the groups' initial ability. The results showed that experimental group scores in all terms were superior to the control group scores. In other words, students in the experimental group could speak English more fluently from the beginning. The result was due to the fact that students of the experimental group and control group attended different classes.

Table 1: Mixed Two-way Repeated Measures ANOVA.

| Measures | | Source of Variation | Sum of Squares | DOF | Mean Square | *F*-value | *p*-value |
|---|---|---|---|---|---|---|---|
| Complexity | Structural complexity | Between learning methods | 11.340 | 1 | 11.34 | 3.288 | 0.079+ |
| | | Among test terms | 6.059 | 1 | 6.059 | 3.436 | 0.073+ |
| | | Interaction | 6.952 | 1 | 6.952 | 3.942 | 0.055+ |
| | Lexical complexity | Between learning methods | 260.521 | 1 | 260.521 | 2.516 | 0.122 |
| | | Among test terms | 96.755 | 1 | 96.755 | 2.137 | 0.153 |
| | | Interaction | 6.209 | 1 | 6.209 | 0.137 | 0.713 |
| Fluency | Speech rate | Between learning methods | 1285.808 | 1 | 1285.808 | 4.454 | 0.0423* |
| | | Among test terms | 1030.779 | 1 | 1030.779 | 25.764 | 0.000**** |
| | | Interaction | 75.034 | 1 | 75.034 | 1.875 | 0.18 |
| | Repair fluency | Between learning methods | 227.390 | 1 | 227.390 | 4.430 | 0.0428* |
| | | Among test terms | 3.709 | 1 | 3.709 | 0.267 | 0.690 |
| | | Interaction | 4.690 | 1 | 4.690 | 0.337 | 0.565 |
| | Silence | Between learning methods | 10.350 | 1 | 10.350 | 0.091 | 0.765 |
| | | Among test terms | 1211.925 | 1 | 1211.925 | 60.528 | 0.000**** |
| | | Interaction | 57.407 | 1 | 57.407 | 2.867 | 0.0996+ |
| Accuracy | | Between learning methods | 0.000547 | 1 | 0.000547 | 0.017 | 0.898 |
| | | Among test terms | 0.000263 | 1 | 0.000263 | 0.011 | 0.918 |
| | | Interaction | 0.0264 | 1 | 0.0264 | 1.078 | 0.307 |

### *5.3 Results in Accuracy*

For Accuracy, there were no significant differences.

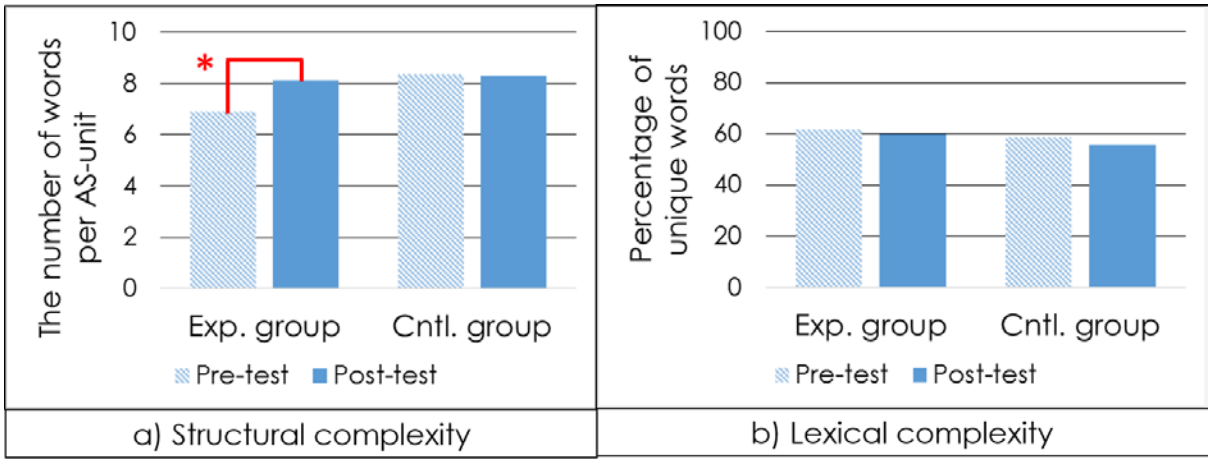


Figure 4. Results for Complexity.

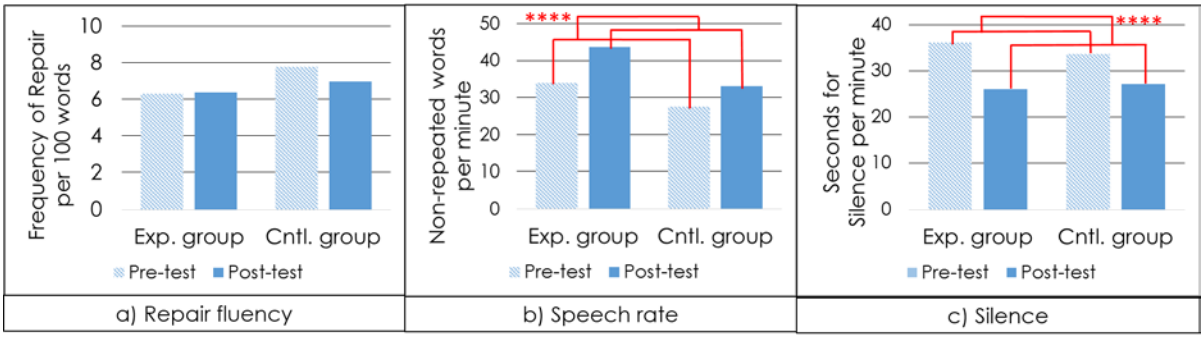


Figure 5. Results for Fluency.

Table 2: The number of reflection on the worksheet (N = 21).

| | Week | | | |
|---|---|---|---|---|
| Reflection | 1st | 2nd | 3rd | 4th |
| Other words, phrases | 8 | 8 | 11 | 5 |
| Added more information | 10 | 12 | 4 | 11 |
| The peer could understand the speech completely | 0 | 1 | 1 | 0 |
| Grammar correction | 1 | 0 | 2 | 0 |
| Improve body language or intonation | 2 | 2 | 1 | 3 |
| Nothing | 2 | 0 | 4 | 3 |

### *5.4 Results of Descriptions on Reflection Worksheets*

Table 2 shows learners' reflections for each week. In reflections, many learners could write detailed improvements using other words or phrases, or adding more information. The result might show that learners could consider more options for expressing what they wanted to say so that the peer could understand more easily.

### *5.5 Discussion*

The learning design could improve Structural complexity but not Repair fluency. The reasons of the results are discussed in the following paragraphs.

In Structural complexity, learners who conducted self-study with scaffolding of MAST and pair reflection could use more words in one AS-Unit, as Figure 4-a illustrates. In other words, the experimental group learners tended to add more words or phrases to explain what they wanted to say. We provide typical examples in pre- and post-tests and explain details in the next paragraph.

In the pre-test, Learner A of the experimental group spoke in very short sentences, "…many people, many people have a little time at lunch time. So, we can eat Ramen for short time. So it is good for lunch to eat." In the post-test, Learner A spoke in this way, "…I like to, I like to take a train and go to another, another place, especially most place Tohoku and Hokkaido." Learner A explained her or his hobby by adding one more phrase. Another example shows that Learner B spoke in this way in the pre-test, "…everyone practiced hard for this. And I practiced hard, too, for it," but this way in the post-test, "The thing that I, I feel the most enjoy is, hmm, entry the contest of the brass band." Learners A and B used more than one verb and added an adverb phrase, an adjective clause, or a noun clause, so that the number of words in one AS-Unit increased.

This effectiveness might show that the reflection activity could facilitate learners in gaining strategies of how to think up and use approximate phrases when they do not retrieve appropriate words to express what they want to say. We can explain it by referring to Kolb's (1994) experiential learning model. Learners in the experimental group spoke the summary in English, and then they reflected on their speech with their peers. During the reflection activity, many learners discussed better expression that enabled their peers to understand their speech, as seen in Table 2. According to Kolb's experiential learning model, when learners have concrete experience and reflect on the experience from a different perspective, they can construct more abstract knowledge. In other words, reflection with a peer stretches the retrieval of words from Lexicon.

MAST scaffolding for summary speaking also has an important role in gaining such strategies. Mackey (1999) observed that in language learning, feedback's effectiveness depends on learners' developmental readiness. Feedback from peers or instructors does not work if the learner does not have appropriate readiness. In this research, learners who obtained MAST scaffolding during self-study might have activated knowledge about vocabulary needed for summary speaking because the scaffolding facilitated learners' use of recommended words in the summary through Short question and answer practice. This might have helped develop readiness, which led to greater effectiveness at stretching learners' knowledge of how to retrieve words from Lexicon.

The unchanged Repair fluency of the experimental group, which differs from the previous result using only MAST (Nakaya and Murota, 2016a), might have been caused by the higher Structural complexity scores. As mentioned previously, experimental group learners might have learned how to access Lexicon from another perspective. Thus, they might have tried to search and retrieve better lexical items or phrases for additional explanation to express what they wanted to say, and sometimes repeat the same words when searching the words. In the post-test, Learner A spoke in this way, "…and go to another, another place, especially most place Tohoku and Hokkaido," which we showed the previous paragraph. During repeating the word "another", the learner might try to search a phrase to add more information, which is "especially most place Tohoku and Hokkaido." This kind of utterance resulted in keeping Repair fluency scores, as shown in Figure 5-a. In addition, although there were no significant differences between pre- and post-test scores of the control group, Repair fluency scores of the control group seemed to improve, which might be due to repeating the summary-speaking task.

## 6. Conclusion and Future Works

In this research, we investigated the effectiveness of the learning design that combines scaffolded summary-speaking self-study using MAST and pair reflection in an English class. After analysis, we concluded that the learning design was effective for improving Structural complexity, Speech rate, and Silence scores. The current learning design compares favorably with the previous design, which was effective only for increasing Speech rate and decreasing Silence and which offered only summary-speaking tasks and speaking to a peer.

From these results, reflection on a learner's speech with a peer might enable the learner to stretch knowledge of how to access and choose vocabulary words because, during reflection activities, learners received feedback from a different perspective for expressing what they wanted to say. In addition, scaffolding practices in self-study play an important role in gaining strategies because scaffolding might activate related knowledge for expression as a pre-task.

In general, future studies need greater clarifying investigation on the effectiveness of reflection activities and scaffolding practices separately. Moreover, we can seek out broader computerized possibilities for assisting language learning, with ICT playing some of the roles mentioned previously.

For example, technologies such as voice recognition and corpus studies might offer functions to develop readiness for interaction with other people by judging learners' achievement of oral tasks and recommending more appropriate phrases during self-study using ICT. Such practice activities might make EFL learners' experiences in face-to-face communication more effective.

## Acknowledgements

We would like to thank Prof. Masatoshi Tamura and Prof. Takehiko Tanioka of Institute for Liberal Arts, Tokyo Institute of Technology for their help in conducting the experiments.

## References

Abrams, Z. I. (2003). The Effect of Synchronous and Asynchronous CMC on Oral Performance in German, The Modern Language Journals, 87(2), 820-833.

Cheng, S. C., Hwang, W. Y., Wu, S. Y., Shadiev, R., & Xie, C. H. (2010). A Mobile Device and Online System with Contextual Familiarity and its Effects on English Learning on Campus, Journal of Educational Technology & Society, 13(3), 93-109.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring Spoken Language: A Unit for All Reasons, *Applied Linguistics*, 21(3), 354-375.

Hardison, D. M. (2004). Generalization of Computer-assisted Prosody Training: Quantitative and Qualitative Findings, Language Learning & Technology, 8(1), 34-52.

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning, Computer Assisted Language Learning, 28(1), 81-96.

Kolb, D. A. (1984). Experiential Learning: Experience as the Source of Learning and Development, Englewood Cliffs, New Jersey: Prentice Hall.

Levelt, W. M. J. (1989). Speaking: From Intention to Articulation, Cambridge, MA: The MIT Press.

Mackey, A. (1999). INPUT, INTERACTION, AND SECOND LANGUAGE DEVELOPMENT: An Empirical Study of Question Formation in ESL, *Studies in Second Language Acquisition*, 21(4), 557-587.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. Studies in Second Language Acquisition, 20, 83-108.

Nakaya, K. & Murota, M. (2016a). Development of a Mobile Application for Listening and Scaffolded Autonomous Summary Speaking and Its Effectiveness for Increasing Fluency, *The Journal of Information and Systems in Education*, 15(1), 48-61.

Nakaya, K. & Murota, M. (2016b). Improvements to a Learning Design: The Combination of Mobile Application for Summary Speaking Task by Self-study and Pair Work in a Class -The effectiveness in scaffolding tasks and visualizing learning achievement-, *The 16th IEEE International Conference on Advanced Learning Technologies - ICALT2016*, 337-341.

Nakaya, K. & Murota, M. (2015). Learning Design in Combination of Mobile Application for Summary Speaking Task by Self-study and Pair Work in a Class, *Work-in-Progress Poster (WIPP) Proceedings of the 23th International Conference on Computers in Education (ICCE) 2015*, pp.22-24.

Olson, D. J. (2014). Benefits of Visual Feedback on Segmental Production in the L2 Classroom, Language Learning & Technology, 18(3), 173-192.

Payne, J. S., and Ross, B. M. (2005). Synchronous CMC, Working Memory, and L2 oral proficiency development, Language Learning and Technologies, 9(3), 35-54.

Skehan, P. (2009). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis, *Applied Linguistics*, 30(4), 510-532.

Swain, M. (2005). The Output Hypothesis: Theory and Research, in Kinkel, R (Eds), Handbook of Research in Second Language Teaching and Learning (pp. 471-483). New York: Routledge.

Tavakoli, P. & Skehan, P. (2005). Strategic Planning, Task Structure, and Performance Testing, in Ellis, R. (Eds), Planning and Task Performance in a Second Language (pp. 239-273). Amsterdam: John Benjamins Publishing Company.

Yanguas, I. (2010). Oral Computer-Mediated Interaction Between L2 Learners: It's about Time! Language Learning & Technology, 14(3), 72-93.