Oral Output Training in EFL Context: The Effects of Task Repetition on Speech Complexity, Accuracy, and Fluency

Akira Iwata

(Hokkaido Musashi Women's Junior College)

Abstract

This study investigates the impact of oral task repetition as a method of EFL classroom instruction on immediate complexity, accuracy, and fluency (CAF) development, and learners' perception of different types of topic introduction. In this study, 20 college students experienced three different types of topics in their oral output training for a six-week period as part of regular class time, in which they orally produced three speeches (of 2 minutes, 1.5 minutes, and 1 minute length, respectively). Topics based on Impromptu, Writing, and Story were each introduced twice. They recorded a retelling of their story in the sixth week to see the immediate effects of the activity. A questionnaire survey was also conducted on their perception of the activity and the types of topic introduction. The results indicated that task repetition significantly improved fluency, but neither complexity nor accuracy improved immediately, and most participants felt Writing topics were easiest.

1. Introduction

1.1 Background

Generally, the objective of language learning is to be able to utilize the target language with enough competence to communicate effectively with others. When we think about language pedagogy at school, especially in an English as a foreign language (EFL) context, like Japan, the objective should be on the improvement of all four macro-language skills, namely, reading, listening, speaking, and writing, with respect to complexity, accuracy, and fluency.

In Japanese high school EFL classroom instruction, it has been repeatedly reported that the traditional grammar-translation method appears to still be in use and not many opportunities to speak and write English are provided. Therefore, Japanese EFL learners are weak in their productive skills, namely speaking and writing. Also, most students have scored around Grade 3 on the Eiken Test in Practical English Proficiency, which is equivalent to A1 (Beginner) level in the Common European Framework of Reference for Languages (CEFR). This is far lower than the Government's goal of 50% of high school graduates scoring Grade 2 or pre-2 on the test (MEXT, 2015).

Specifically, it is widely recognized that opportunities for oral output are limited compared to writing, partly because of university entrance examinations that emphasize writing more than speaking. Indeed, Japanese EFL learners are weak at speaking compared to the other skills (Matsuda & Gobel, 2004). Furthermore, many Japanese students studying abroad experience learning difficulties and lack confidence in speaking due to their past learning experiences, which emphasized

grammar and reading skills rather than conversational skills in teacher-centered classrooms (Sawir, 2005). This means that providing more opportunities for oral practice should be prioritized before sending students for their tertiary education abroad.

Many high school and college students reported being insecure about their speaking skills. However, based on the author's research, among the four macro-language skills, students sought to improve their speaking skills the most (Iwata & Suzuki, 2017). They unanimously reiterate that they wanted to be fluent in English, which generally meant being able to speak reasonably fast with comprehensible output for the interlocutor. Given this situation, there is potentially a high demand from Japanese EFL learners to improve their speaking skills. Hence, feasible and effective classroom practices that improve learners' speaking skills, including complexity, accuracy, and fluency (CAF), are imperative, and if we focus on Japanese learners' needs, fluency should be prioritized in instruction, as Nation (2013) suggests.

1.2 Task Repetition for CAF Development

A considerable amount of research has been conducted into the effects of task repetition on oral output. First, with regard to the nature of speaking, one of the most influential models is Levelt's (1989) speech production model (Bygate, 2005). This model consists of three stages that speakers go through, namely, conceptualization, formulation, and articulation. The first stage, conceptualization, concerns the selection and retrieval of relevant information to formulate intended meaning. In the formulation stage, preverbal messages are converted into linguistic structures, and in the articulation stage, speakers

verbalize the output. Task repetition affects performance because of the following mechanism. In the first session of speech delivery, the speakers focus on establishing the message content. However, in the next session, they are already familiar with the speech content; hence, they will focus more on the selection and monitoring of the language, which, accordingly, improves the language's complexity, accuracy, and fluency (Bygate, 1999).

Furthermore, in designing classroom practices in EFL context, three important components must be considered, namely, pre-task planning, noticing as a linguistic factor, and anxiety as an affective factor.

Planning is crucial for producing output, and almost any type of production includes some kind of planning (Ellis, 2005). Pre-task planning in this research is defined as providing enough time for learners to work mentally on their oral output before actually performing the task (Ahmadian, 2012). Unlike online planning, which promotes accuracy because of more attention to form, pre-task planning allows more attention to message conveyance and results in greater fluency and lexical variety (Yuan & Ellis, 2003). Skehan's Limited Capacity Hypothesis (1998) predicts trade-off effects among complexity, accuracy, and fluency because of limited mental resources and working memory. When speakers perform an oral task, they are not able to allocate equal attention to those three dimensions of performance; therefore, attending to one of them might result in lower performance in the others (Skehan, 2009). Although Robinson (2005) argues in his Cognition Hypothesis that a cognitively and functionally demanding task will encourage learners to produce more complex and accurate output, some

studies that used oral tasks, unlike written task, did not support it (e.g. Robinson, 2007). Also, it is considered that accuracy is more susceptible to individual language ability than cognitive task difficulty (Sasayama, 2011). Native speakers can deal with multiple demanding tasks with ease because their language knowledge has been proceduralized; however, since this study focused on oral task and non-native speakers of English, Skehan's hypothesis is considered to be more compatible. Wigglesworth (1997) reported that even pre-task planning for one minute helped speakers improve their production, and resulted in greater fluency. Japanese EFL learners can deliver a speech or presentation reasonably well if they are allowed to prepare the content of the speech and rehearse in advance (Bygate, 1999). Therefore, pre-task planning should be considered an important component of speaking instruction in Japanese classrooms.

Noticing plays a crucial role in second language acquisition, as evidenced in Swain's output hypothesis (1993). In the context of oral task repetition, it reported that accuracy, as well as fluency, improved when participants were given opportunities to attend to forms after the initial speech (Date & Takatsuka, 2012; 2013). In their study, the participants were required to perform a narrative task with a six-picture strip cartoon. In each session, the experimental group transcribed their first narrative and self-monitored it. Next, a native speaker checked it without explicit correction in order to provide them the chance to notice erroneous forms. Finally, explicit and direct feedback was provided before they read it three times silently and told the story again. The control group narrated the story without any feedback. The results of the posttests suggested both groups did

not improve in accuracy and fluency when they tackled different strips from the pretest but both groups improved in accuracy and fluency when they tried the same strips, and the group that was allowed to notice forms performed better. They concluded that proceduralization of linguistic knowledge leads to greater accuracy and fluency in the same task and greater accuracy in the new task. However, it could be too much burden for students to make a story from strips, and the quality could vary according to individual ability. Baleghizadeh and Derakhshesh (2012) reported that, even after the speaker was fully aware of their erroneous forms, they could not remove several types of errors, such as third person singular, prepositions, and verb tenses, from the subsequent performance. Furthermore, it is significantly more difficult to self-repair mistakes when the learners' proficiency levels are insufficient. It is also very hard to provide immediate, individual feedback to all the participants when they consecutively repeat the task in the same class period. For these reasons, to be more realistic to the classroom setting, they were only allowed to mentally reflect on their speech, both on its content and forms using a set of six pictures. No corrective feedback was given on the participants' speeches.

Anxiety is believed to affect the development of second language proficiency (Housen, Kuiken, & Vedder, 2012). In speaking, two types of anxiety were highlighted by Hiromori (2014), which relate to (a) message generation and (b) transformation of the message into a particular linguistic form, which pertain to the conceptualization and formulation stages of Levelt's model (1989), respectively. In a real-life situation, speakers would often be under pressure with little planning time, thus resulting in anxiety. However, considering the EFL

language class context, especially where learners' proficiency levels are low, it is not appropriate to put learners under too much pressure. A sense of security should be provided for learners with high anxiety, low risk-taking tendencies, low confidence, and low motivation (Yousefi, 2016). Therefore, learners should be provided with hints on what to say (message generation), and how to say it (transformation of the message into a particular linguistic form) by presenting visuals (some pictures showing the contents of the story).

This study adopted pre-task planning, visual aids, and allowed very short amounts of time for attending to language forms. Also, few studies have reported learners' perceptions of the types of topic introduction, this study adopted three different types of topic introduction (Impromptu topics concerning personal experiences and beliefs, Writing topics, and Story topics) to investigate the participants' perspectives on difficulty and enjoyment.

1.3 Study Purpose

The primary aims of this study are the following two points. One is to determine whether a particular type of oral task repetition in EFL settings can improve learners' immediate speech in respect to fluency, as well as complexity and accuracy, over the course of one trial (1 week). The other is to determine their perception of the three different types of topic introduction (3 topics x 2 weeks). The results would contribute to EFL classroom pedagogy, especially in Japanese secondary and tertiary education.

2. Method
2.1 Participants

This study was conducted in a women's junior college in Japan. All 20 students, whose age range was 18 to 20, enrolled in a Writing II course agreed to participate. They were all English majors and their self-reported English proficiency levels were somewhere between Grade pre-2 to Grade 2 in the Eiken Test in Practical English Proficiency or around the A2 to B1 in CEFR. Due to absences, only 16 complete datasets were available for analysis. All the tasks given were part of the students' regular class requirements. Their first language was Japanese, and none of them had ever stayed outside of Japan for more than two weeks. There was no control group in this research because it used an intact class unit and the participants were assigned to the class based on the results of an institutional placement test. Therefore, there were no other class units whose proficiency was equivalent and the classroom instructional methodology comparable.

2.2 Material

2.2.1 Classes

The instruction took place for six out of fifteen weeks. Each class lasted 90 minutes, and the first 20 minutes were allocated to the speaking activities and the remaining class period comprised of writing about a particular topic, or grammar instruction using a textbook. The first five weeks were spent investigating the participants' perceptions of the different topic types, and the final week was spent researching the CAF development in their English speeches, as well as their perceptions of the different topics.

2.2.2 Speaking Topics

Three different types of topic introduction were adopted for this activity. The first one was (1) Impromptu, in which the

topics were introduced just before the activity began and their contents were relevant to their lives, which were as follows: "My hobbies," and "What I did yesterday." They were allowed to choose no more than six words as cues for their oral presentation. The second type was (2) Writing, in which the topics were chosen from the topics they had attempted as writing topics for assignments outside the class period and were submitted during this Writing II course. "The person who has greatly influenced me" and "What I like about myself" were chosen as speaking topics so that they could speak based on their memory. Again, they were permitted to choose only six words as cues for the talk. The third type was (3) Story, in which they retold a story they had just read assisted by the use of six picture illustrations. The group of first speakers and second speakers were given different stories. Before the actual talk, they were allowed to check their understanding of the story with peers who were given the same story. The stories were taken from a book by Heyer (2004). The participants experienced each topic introduction type twice as shown in Table 1. The training procedure is shown in Table 2.

Table 1
Description of the Training Procedures (Week 1-5)

	1st speakers	2nd speakers
3min	planning	planning
2min	1st speech	listening
1.5 min	2nd speech	listening
1 min	3rd speech	listening
2min	listening	1st speech
1.5 min	listening	2nd speech
1 min	listening	3rd speech

Table 2

Description of the Training Sessions

Week	Topics	Category	Interlocutor
Week 1	My hobbies	Impromptu	peers
Week 2	The person who has greatly influenced me	Writing	peers
Week 3	Speed/ The best doctor	Story	peers
Week 4	What I did yesterday	Impromptu	peers
Week 5	What I like about myself	Writing	peers
Week 6	Hawaiian vacation	Story	PC recording

2.2.3 Questionnaires and Interviews

A questionnaire asking the participants the following three points: (1) how they felt about their performance on the sixth activity, (2) which type of topic introduction that they felt was easier, and (3) which type of topic introduction that they felt allowed for the most enjoyment was distributed. A follow-up interview was conducted a week after the treatment in order to clarify and confirm their responses.

2.3 Procedure

2.3.1 Training Sessions

Before the sixth week, the participants underwent five weeks of training sessions. In the first week, the instructor introduced the aim of the activity and the procedures to follow. The participants were seated in four columns, and the instructor assigned two moving columns for the first speakers, who would only move and change seats in the designated direction, in order to be paired with a different listener each time. Approximately 3 minutes were allocated for planning, after which the speaker spoke on the topic for 2 minutes, then moved on and spoke for 1.5

minutes to a different listener, and finally moved again to talk to another listener for 1 minute. Then, they switched roles, so that the previous listeners were the speakers, and the procedure was repeated again. Each talk was performed relatively soon after the previous one. Therefore, there was only approximately 1 minute for the participants to review their oral production on content, grammar, or vocabulary during the session, including the time they changed partners.

2.3.2 Data Collection Sessions

In the sixth week, the class was conducted in a computer room and each participant was assigned to one computer to record their talks. Unlike the third week, all the participants were given the same story, namely "Hawaiian Vacation," to retell. This story was chosen for data collection because they have to retell what they just read. However, the other categories, Writing and Impromptu, could be spoken based on their experience. Therefore, personal experience could affect their performance. In order to minimize the influence of different personal experiences, Story was chosen for data collection.

For preparation, they read the story and talked with their peers to confirm their understanding. Then, they used headsets to record their retelling. They clicked the start button before retelling the story for 2 minutes after which the instructor would cue them to stop, and they clicked the stop button. In the second and third delivery, they repeated the same procedure for 1.5 minutes and 1 minute, respectively. Approximately one minute was given to the participants for preparing for the next trial. Immediately after the audio recording, they were asked to submit the sound files to the instructor. A questionnaire survey was conducted after the session, and a follow-up interview took

place the following week.

2.4 Data Analysis

For the first source of data, all speeches were initially transcribed verbatim by the author. Then, the transcription was revised by eliminating words and phrases that were repeated, reformulated, or replaced (Foster, Tonkyn, & Wigglesworth, 2000).

Following Ahmadian (2011), complexity was measured as overall complexity, which is the mean length of AS-units (Analysis of Speech units) in a learner's speech, meaning the average number of words in an AS-unit. AS-unit is defined as a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses associated with either of them, and the AS-unit was deemed appropriate because it is essentially a syntactic unit that is a genuine unit of planning (Foster, Tonkyn, & Wigglesworth, 2000). Following Ahmadian and Tavakoli (2010), accuracy was measured by the rate of error-free clauses and the rate of correct verb forms. The former was obtained by dividing the number of correct clauses, in terms of syntax, morphology, and lexicon, by the total number of clauses. The latter was obtained by dividing the number of verbs that are used correctly, in terms of tense, aspect, modality, and subject-verb agreement, by the total number of verbs used in the speech. Fluency was measured by the pruned words per minutes, which was obtained by counting the number of words in the revised transcriptions and calculating words per minute (WPM), as in Foster, Tonkyn, and Wigglesworth (2000). They defined the term "pruned words" as the total number of words, excluding self-repaired words or repetitive words.

The second data source was the results of the questionnaire. The third, and final, source of data was the students' follow-up interviews, where students were asked about what they had written on the questionnaire for purposes of confirmation and clarification. Friedman tests were conducted for complexity (mean number of words for AS-unit (First delivery/Second delivery/Third delivery), accuracy (rate of error-free clauses (First delivery/Second delivery/Third delivery), rate of correct verb forms (First delivery/Second delivery/Third delivery), and fluency (WPM (First delivery/Second delivery/Third delivery)) as the within-subjects variable, in order to determine the effects of task repetition on CAF development. In the case of significant variance among the deliveries, the Scheffé's method test was adopted. Additionally, chi-square tests were conducted for the participants' responses to the questions on the different types of speaking topics. Ryan's method was also adopted in the case of significant variance among participants' preferences.
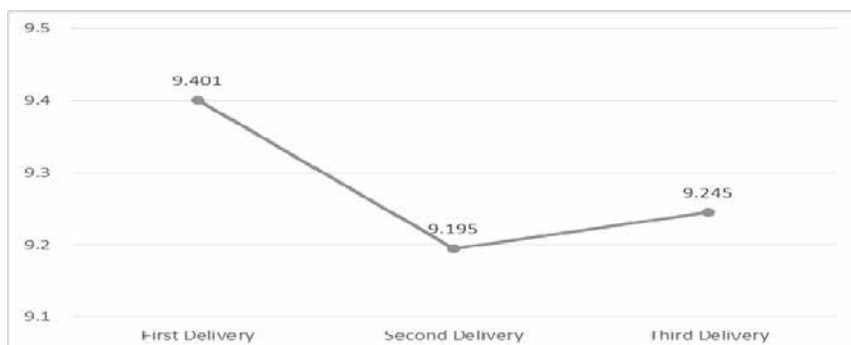
3. Results

3.1 Complexity (Overall Complexity)

Table 3 shows the descriptive statistics of the overall complexity (mean number of words per AS-unit of the deliveries (First, Second, and Third). The Friedman's test revealed no significant difference ($\chi2$ (2) = 1.625, p = .444, Cramer's V= .23) with small effect sizes (see Figure 1).

Table 3
Means and Standard Deviations of Overall Complexity

Delivery	Mean	SD	n
First Delivery	9.401	2.182	16
Second Delivery	9.195	2.029	16
Third Delivery	9.245	1.373	16

Figure 1. Overall Complexity of Deliveries.

3.2 Accuracy

3.2.1 Error-free Clauses Rate

Table 4 shows the descriptive statistics of the rate of error-free clauses of the deliveries (First, Second, and Third). The Friedman's test revealed no significant difference ($\chi 2$ (2) = 0.644, p = .725, Cramer's V= .14) with small effect sizes (see Figure 2).

Table 4
Means and Standard Deviations of Error-free Clauses Rate

Delivery	Mean	SD	n
First Delivery	0.468	0.192	16
Second Delivery	0.499	0.198	16
Third Delivery	0.512	0.171	16

Figure 2. Error-free Clauses Rate of Deliveries.

3.2.2 Correct Verb Forms Rate

Table 5 shows the descriptive statistics of the rate of correct verb forms of the deliveries (First, Second, and Third). The Friedman's test revealed no significant difference ($\chi2$ (2) = 2.0, *p* = .368, Cramer's *V*= .25) with small effect sizes (see Figure 3).

Table 5
Means and Standard Deviations of Correct Verb Froms Rate

Delivery	Mean	SD	n
First Delivery	0.668	0.170	16
Second Delivery	0.710	0.195	16
Third Delivery	0.710	0.190	16

Figure 3. Correct Verb Forms Rate of Deliveries.

3.4 Fluency (Pruned WPM)

Table 6 shows the descriptive statistics of the WPM of the deliveries (First, Second, and Third). The Friedman's test revealed significant main effects of repetition ($x2$ (2) = 28, p < .001, Cramer's V= .94) with large effect sizes. Given the significant difference between deliveries, Scheffé's method test was conducted as post-hoc multiple comparisons. The results of analysis confirmed that there was a significant difference between the First and the Second Delivery ($x2$ = 4.5, p < .001), between the Second and the Third Delivery ($x2$ = 28.1, p < .001), and between the First and the Third Delivery ($x2$ = 10.1, p < .001) (see Figure 4).

Table 6
Means and Standard Deviations of Words per Minute (WPM)

Delivery	*Mean*	*SD*	*n*
First Delivery	62.219	13.897	16
Second Delivery	72.837	15.679	16
Third Delivery	101.062	16.668	16

Figure 4. Words per Minute (WPM) of Deliveries.

3.5 Questionnaires and Interviews

A chi-square test was performed to determine whether the participants felt that the three topic introduction types were equally preferable or challenging. The results indicated a significant difference in the degree of challenge, ($x2$ (2, n = 16) = 6.126, p < .05), but no significant difference in the degree of enjoyment ($x2$ (2, n = 16), = 3.875, *n.s.*). The Ryan's method as post-hoc multiple comparisons revealed that a significant difference was found between Writing and Impromptu on the level of challenge, (p = .026) (see Table 7).

Table 7

The Questionnaire Result 1

Item (n = 16)	Impromptu	Writing	Story	$x2$
Which type of topics did you find easiest?	6% (1)	56% (9)	38% (6)	s.
Which type of topics did you find the most enjoyable?	19% (3)	25% (4)	56% (9)	n.s.

Table 8 presents the results of the questionnaires using Likert scales for the participants' perception of their performance and of the speaking activity itself.

Table 8

The Questionnaire Result 2

Item (*n* = 16)	Yes / 5	4	3	2	No / 1
I was able to speak faster on the first delivery than the second delivery	25%	50%	19%	6%	0%
I was able to speak faster on the second delivery than the third delivery	19%	50%	13%	19%	0%
I was able to speak with more accuracy on the second delivery than the first delivery	6%	38%	44%	13%	0%
I was able to speak with more accuracy on the third delivery than the second delivery	19%	31%	31%	19%	0%
Did you reflect on the content of the previous speech and modify it the next time? Why?	81%				19%
Did you reflect on the grammar of the previous speech and modify it the next time? Why?	63%				38%
Did you enjoy the speaking activity?	19%	63%	19%	0%	0%

Likert Scale: 5 strongly agree, 4 agree, 3 neutral, 2 disagree, 1 strongly disagree

The following results summarize typical responses to the follow-up interviews. For the level of challenge of the topics, many responded that Writing was the easiest because they could still remember the content of their writing. Some participants who chose Story said that the stories were easy to recall because they were interesting and had a flow; also, the use of pictures was considered to be especially beneficial.

On the most enjoyable type of topic, the participants who chose Story said that their partners' stories differed from theirs, and the stories were interesting. Additionally, they listened to the stories three times, which increased their understanding. Another interesting response from participants who chose Impromptu or Writing was that the activity was useful for them in finding out what their peers were thinking, as the contents of the speeches differed from person to person.

As for the last storytelling activity, which was recorded, the

typical answer to the questionnaire was that they felt they could speak faster as they repeated their speech. However, they were not confident that their grammar accuracy had improved in the subsequent production. They also responded that they tried to review and improve the content and grammar, but they could not do so because of time constraints. The responses are discussed in more detail in the next section.

4. Discussion

One of the purposes of the current study was to determine whether a particular type of oral task repetition affects a learner's immediate oral performance in terms of CAF. The results indicated that only fluency was significantly improved. The participants' perceptions of the improvement of oral fluency were concurrent with the actual results. Importantly, 75% and 69% of the participants felt that they were actually more fluent in the second and third deliveries, respectively, thereby illustrating the benefit of the activity, as a feeling of improvement would most certainly affect motivation (Dörnyei, 2001). Indeed, most of the students stated that they enjoyed the activity because they felt they could speak faster, which was a feeling they rarely experienced previously. Another noteworthy aspect is the different conditions during data collection when the participants recorded their speeches. As they were exempt from the role of listener, they could focus on their role as speaker. With no audience (listeners) in front of them, the participants felt it was easier to speak.

Fluency development is mainly attributed to three factors: planning, selective attention, and repetition. Ellis (2005) argued that even the most carefree speech requires some degree of

planning, and in this study, there was no time pressure associated with planning. Therefore, participants could check the content of their speeches by confirming the facts and flow of the story with their peers. Additionally, the use of pictures further guided their speeches. These conditions could have lessened their mental load in the pre-task planning, and participants could familiarize themselves with the content, and focus on the use of language, which resulted in improved fluency (Bygate, 1999) as well as enhanced word selection, morphemes, and grammatical structures (e.g., Bygate, 2001). This study's results support previous research that dealt with the impact of pre-task planning on fluency (e.g., Sangarun, 2005; Yuan & Ellis, 2003).

In Skehan's (1998) Limited Capacity Hypothesis, because of limited mental resources, the trade-off effects may result in a lower performance in either language complexity, accuracy, or fluency. In this study, because the participants knew that they had to speak three times, with the time limit being reduced each time, they might have prioritized fluency because, based on Ellis and Barkhuizen (2005), the type of task or instruction can have the result of encouraging learners to prioritize one component over the others, which could have influenced the results. Regarding the participants' perception of accuracy, 44% and 50% felt that they spoke more accurately in the second and third trial, respectively. This coincides with the author's deduction from the comparison of the transcriptions. In fact, the percentages of error-free clauses and correct verb forms showed a slight improvement, though this was not significant. Participants had a very short time to reflect on their speech between the trials. However, they were not allowed to confirm the original written

story or receive feedback on their speech. Therefore, all they could do was recall the story through the pictures and self-monitor their speech. If they were given enough time to reflect on their speech using the original written story, and somehow given feedback on their speech, the results might be different. In fact, 81% and 63% of them reported that they tried to review and modify the content or grammar in the second and third trial respectively. However, in the interview, many responded that they could not afford to modify the content and grammar because of time limitations; thus, they prioritized finishing the retelling of the story instead. As for grammar, several students indicated it was a low priority in their speech. This could be interpreted as they simplified the use of vocabulary, grammar, and sentence structure because they did not have enough time to modify them online. In terms of complexity, the second and third trial scored lower than the first, though they were not significant. Because the participants prioritized fluency, meaning that they tried to finish telling the story within the limited time, it could be possible that the AS-unit tended to be short. As in the case of accuracy, the trade-off effects could have affected the results.

Immediate repetition improves fluency (Snellings, Van Gelderen, & De Glopper, 2004), most likely because the repeated use of vocabulary and grammatical constructions can facilitate their retrieval through lexical and syntactic priming (Youjin & MacDonough, 2008). Repeated practice increases the retrieval speed of words and phrases, and induces learners to perceive the words and phrases as formulaic sequences (De Jong & Perfetti, 2011). Hence, the lexical and grammatical structures that are used remain activated for immediate retrieval. From the

comparison of learners' speech transcriptions, each participant used almost the same vocabulary items and grammatical constructions throughout the three trials. Therefore, language complexity and accuracy did not seem to improve much.

The other purpose of this study was to research learners' perceptions of the topic introduction. Regarding the type of topic, the perceived easiest activity was Writing, and the most enjoyable was Story. Only Writing was perceived as being significantly easier than Impromptu. Although no significance was found between it and the other methods, Story was the most enjoyable. In the interviews, many said they enjoyed Story most because the story they listened to differed from the one they had read. In addition, deeper understanding and new knowledge were obtained every time they heard the story, which was exciting for them. They highlighted that the contents of Writing and Impromptu were often somewhat similar and sometimes predictable, which made them perceive them as less exciting. As speakers, some participants claimed that, in Writing and Impromptu, they sometimes had to divulge something personal, which bothered them, but this was not the case in Story. The interviews also revealed that Writing was the easiest because they remembered the content well, as they had revised it in class. Furthermore, Story used pictures as a visual aid, which gave them more information than just words, but Impromptu only used words as supportive cues. It is important that the task assigned should be reasonably challenging and manageable for students. However, if the difficulty levels of the topics are equal, teachers should consider which task is more enjoyable for their students because task enjoyment could also affect learner incentives.

5. Conclusion and Future Research

This study investigated both the effects of task repetition on immediate oral CAF development and learners' perceptions of different types of topic. Task repetition in this research is considered effective in improving learners' oral fluency, and thus it should be adopted in EFL classrooms because of the following points. First, task repetition in this study is effective in improving oral fluency, at least temporarily. Second, as the interview protocols show, learners experience the improvement and achievement in their oral skills, which is significant because it can lead to the development of increased motivation and confidence, especially in the current circumstances of insufficient oral practice in classrooms, where a great deal of learners' needs exist. Finally, the repeated words, phrases, and sentence structures could be learned and ready to use in similar situations or on similar topics in the future because repetition is essential in language learning, especially in the EFL context (Larsen-Freeman, 2009). Future study should focus on what type of topics and tasks can be feasible and effective in improving students' oral outputs in complexity, accuracy, and fluency in EFL classrooms. Also, they should consider the longer term effects of the practice and how it will contribute to practitioners' decision making.

References

Ahmadian, M. J. (2011). The effect of 'massed' task repetitions on complexity, accuracy and fluency: does it transfer to a new task? *The Language Learning Journal, 39*, 269-280.

Ahmadian, M. J. (2012). The effects of guided careful online

planning on complexity, accuracy and fluency in intermediate EFL learners' oral production: The case of English articles. *Language Teaching Research, 16*, 129-149.

Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*, 35-59.

Baleghizadeh, S., & Derakhshesh, A. (2012). The effect of task repetition and noticing on EFL learners' oral output. *International Journal of Instruction, 5*, 141-152.

Bygate, M. (1999). Task as the context for the framing, re-framing and unframing of language. *System, 27*, 33-48.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (PP. 23-48). Harlow, UK: Pearson Longman.

Bygate, M. (2005). Oral second language abilities as expertise. In K. Johnson. (Ed.), *Expertise in Second Language Teaching* (pp. 104-127). London: Palgrave.

Date, M., & Takatsuka, S. (2012). Effects of task repetition on oral performance of the same task and a new task of the same type. *ARELE, 23*, 345-360.

Date, M., & Takatsuka, S. (2013). The effect of task repetition and noticing of forms on proceduralization of linguistic knowledge. *Kyoiku Jissen Ronshu, 14*, 89-101.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*, 533-568.

Dörnyei, Z. (2001). *Teaching and researching motivation*. Harlow, UK: Pearson Education Limited.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in second language* (pp. 3-34). Amsterdam: John Benjamins.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics, 21*, 354-375.

Heyer, S. (2004). *All new easy true stories.* New York: Longman.

Hiromori, T. (2014). What are the determining factors of L2 improvised speech performance? An exploratory study. In M. Paradowski (Ed.), *Teaching languages off the beaten track* (pp. 111-126). Frankfurt: Peter Lang.

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency*. Amsterdam: John Benjamins.

Iwata, A., & Suzuki, A. (2017). Teacher feedback's effect on EFL writing: A case study on text quantity, incentives, and writing confidence. *Research Bulletin of English Teaching, 13, 57-80.*

Larsen-Freeman, D. (2009). Adjusting expectations: the study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*, 579-589.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*, 21-36.

MEXT. (2015). Abstract of survey on the English ability of third-year high school students. Retrieved from 2016.12.10. http://www.mext.go.jp/component/a_menu/education/detail/__icsFiles/afieldfile/2015/07/03/1358071_01.pdf

Nation, P. (2013). *What every EFL teacher should know?* Tokyo: Compass Publishing.

Robinson, P. (2005). Cognitive complexity and task sequencing: A review studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*, 1-33.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching, 45*, 193-213.

Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111-141). Amsterdam/Philadelphia: John Benjamins.

Sasayama, S. (2011). Cognition hypothesis and second language performance: comparison of written and oral task performance. *Second Language Studies, 29*, 107-129.

Sawir, E. (2005). Language difficulties of international students in Australia: The effects of prior learning experience. *International Education Journal, 6*, 567-580.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*, 510-532.

Snellings, P., Van Gelderen, A., & De Glopper, K. (2004). The effect of enhanced lexical retrieval on second language writing: A classroom experiment. *Applied Psycholinguistics, 25*, 175-200.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158-164.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing, 14*, 85-106.

Youjin, K., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics, 29*, 149.

Yousefi, M. A. (2016). The influence of affective variables on the complexity, accuracy, and fluency in L2 oral production: The contribution of task repetition. *Journal of English Language Teaching and Learning Tabriz University, 17*, 26-48.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, 1-27.