兵庫教育大学　教育実践学論集　第14号　2013年３月　pp.89－101

The Effect of Task Repetition and Noticing of Forms on Proceduralization of Linguistic Knowledge

DATE Masaki *, TAKATSUKA Shigenobu **

（平成24年６月19日受付，平成24年12月６日受理）

タスクの繰り返しと言語形式の気づきが言語知識の手続き化に及ぼす効果

伊　達　正　起*, 髙　塚　成　信**

参加者に絵から話を作る物語タスクを pretest として与え，タスクを繰り返した後に pretest と異なるタスク (posttest 1) と pretest と同じタスク (posttest 2) を与えた。そして，タスクを繰り返すだけのグループ 1 とタスクを繰り返す時に言語形式に気づかせたグループ 2 のタスク遂行時の流暢さ（ポーズの長さ，発話時間全体に対するポーズを除く発話時間の割合，流暢な連続部の長さ）と正確さ（冠詞と動詞のエラー率）を調べた。その結果，posttest 1 では両グループとも練習前に比べ流暢さと正確さは向上せず，流暢さでグループ間に差はないが，正確さでグループ 2 は低下しなかった。posttest 2 では両グループとも練習前より流暢さと正確さが向上し，流暢さでグループ間に差はないが，正確さでグループ 2 が上回った。

キーワード：言語知識の手続き化，タスクの繰り返し，言語形式の気づき

1. Background

In English classes, students are often given opportunities to repeat their speaking. For example, students are given a speaking task and do the task in pairs first, then change partners and do the same task again, repeating the task several times. When we observe how students are speaking, it seems that they become more familiar with and confident in doing the task by repeating the same/similar information to a succession of partners. Consequently, they seem to enjoy speaking and speak more and more. However, when students speak, they mainly focus on what they want to say, i.e., meaning, not on how they speak, i.e., form. Furthermore, the students tend simply to repeat the speaking task and not give feedback to each other. Therefore, the following questions arise:

(1) Are students, who seem to enjoy speaking and speak more and more through task repetition, actually speaking with increased fluency?

(2) Are students, who seem to enjoy speaking and speak more and more through task repetition, actually speaking with increased accuracy?

Although task repetition has been used in classrooms for several decades, few studies had empirically investigated its effects. However, in recent years, several studies have started to focus on the effectiveness of task repetition and have shown its effectiveness in the second performance. For example, compared to the first performance, accuracy improved (Lynch & Maclean, 2001[(1)]), both fluency and complexity improved (Ahmadian & Tavakoli, 2011[(2)]; Bygate, 2001[(3)]), or both accuracy and complexity improved (Gass, Mackey, Fernandez, & Alvarez-Torres, 1999[(4)]). A possible reason for this consistent improvement in the second performance may be that suggested by Bygate (1999)[(5)]: students are likely to focus initially on message content, then, subsequently, once the message content and the basic language needed to encode it have been established in the first performance, to switch their attention to the selection and monitoring of appropriate language in the second performance. Thus, these results may indicate that students switch from their focus on message content to attention

* 兵庫教育大学大学院連合学校教育学研究科学生（Doctoral program student of the Joint Graduate School in Science of School Education, Hyogo University of Teacher Education）

** 岡山大学大学院教育学研究科（Graduate School of Education, Okayama University）

to language through task repetition. However, these studies do not show precisely what cognitive processes are at work to improve the second performance; that is, if "there is some change in the learner's L2 knowledge representation" (Ellis, 2005[6], p. 27).

Accordingly, De Jong and Perfetti (2011)[7] tried to investigate the influence of task repetition on 'proceduralization' defined as "changes in underlying cognitive mechanisms" (p. 563). Their study was based on ACT-R (Adaptive Control of Theory-Rational) (Anderson & Lebiere, 1998[8]; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004[9]), a skill-acquisition theory in cognitive psychology that defines language learning as one type of general cognitive learning. The learning of ACT-R develops through three stages: declarative, procedural, and automatic. The first stage is where declarative knowledge of the skill or cognitive act (explicit knowledge of what the skill or act is) is acquired; in the second stage, the skill or cognitive act is repeatedly used or performed with the declarative knowledge so that procedural knowledge of the skill or act (implicit knowledge of how to use the skill or act) is stored and develops; and, in the third stage, the skill or cognitive act continues to be used or performed until it is automatized and performed speedily and flawlessly. Procedural knowledge is necessary for language use; therefore, the second stage of storing and developing procedural knowledge is thought to be an indispensable process for language learning. According to ACT-R, this process is called 'proceduralization', whereby can occur both the construction of new production rules (steps of cognition and the basic form of 'goal condition + chunk retrieval ➡ goal transformation') and the collapsing of smaller production rules into larger ones.

De Jong and Perfetti (2011) focused on the effectiveness of task repetition on proceduralization for fluency development. They measured fluency development by using three measures of fluency: the mean length of pauses, the phonation/time ratio (the percentage of time spent speaking as a proportion of the total time taken to produce the speech sample), and the mean length of fluent runs. These measures are good predictors of fluency (Kormos & Dénes, 2004[10]), and, when used in combination, can be indicators of proceduralization (Towell, Hawkins, & Bazergui, 1996[11]). De Jong and Perfetti (2011) used a 4/3/2 task (Nation, 1989[12]), in which students did a speaking task for four minutes and then retold it twice, as close to verbatim as possible, in three and two minutes. Twenty-four students enrolled in speaking courses at a high intermediate level in an institute for ESL at a university performed three 4/3/2 tasks and were given three speaking tests. Monologue tasks were used in the sessions and tests, and the students were given a topic (e.g., *How do you feel about pets? Do many people have pets in your country? How are they treated, in general?*) and spoke about it. For each 4/3/2 task, students in the repetition group did the same task three times whereas students in the no repetition group did a new task three times. In the pretest before the training sessions, the immediate posttest one week after the sessions, and the delayed posttest four weeks after the immediate posttest, all the students did a new task for two minutes. It was then found that both groups increased fluency during a training session. However, only the students in the repetition group maintained this increase in fluency in the two posttests. Therefore, De Jong and Perfetti concluded that the task repetition in the 4/3/2 task may cause proceduralization, and result in an increase in fluency and the long-term retention of the increased fluency, as well as transfer of the increased fluency to a new task.

However, there are three questions left unanswered by the study of De Jong and Perfetti:

(1) It is not clear if proceduralization only facilitates an increase in fluency of a new task.

According to ACT-R (Anderson & Lebiere, 1998), declarative knowledge takes the form of chunks in the declarative module, and procedural knowledge consists of production rules in the production module. ACT-R also claims that each production rule is triggered by a goal in the intentional module and retrieves one or, at most, a few declarative chunks. However, it could be possible that an erroneous chunk is retrieved from the declarative module and used in a mistaken production rule. Furthermore, new production rules can subsequently gain strength so as to be able to compete with previously existing rules through repeated practice (Anderson et al.,

2004). Therefore, it could be possible that, even if a new production rule is erroneous, repeated practice of the rule may facilitate its proceduralization, and fluency may improve. The process of acquisition can affect some changes in the learner's second language knowledge representation in terms of accuracy of the performance in the same task and/or in a new task, and/or fluency of the performance in the same task and/or in a new task (Ellis, 2005). Proceduralization also shows "changes in underlying cognitive mechanisms" (De Jong & Perfetti, 2011, p. 564). In other words, it is necessary to prove the influence of task repetition on proceduralization by not only fluency in a new task but also fluency of the same task and accuracy in the same task and in a new task.

(2) It is unclear if just repeating a task facilitates proceduralization.

According to ACT-R, proceduralization is the process of storing and developing procedural knowledge of the skill or cognitive act (implicit knowledge of how to use the skill or act) stored in the production module. ACT-R also claims that for the facilitation of procuduralization, the skill or act must be repeatedly used or performed with declarative knowledge of the skill or act (explicit knowledge of what the skill or act is) stored in the declarative module. When a learner repeats the first speech immediately, the learner has the benefit of having used certain grammatical constructions, which can facilitate retrieval through syntactic priming (De Jong & Perfetti, 2011; Kim & McDonough, 2008[13]). In addition, when learners repeat a task, they are likely to switch their attention to the selection and monitoring of appropriate language (Bygate, 1999). Therefore, task repetition itself seems to facilitate the skill or act to be repeatedly used or performed with declarative knowledge to some extent. However, it was found that accuracy does not improve without specific instructions about what learners should pay attention to (Bygate, 2001). Furthermore, specific, discrete instructions about what forms learners should pay attention to before undertaking a task can make them more mindful of those forms during the task performance (Mehnert, 1998[14]; Sangarun, 2005[15]; Yuan & Ellis, 2003[16]). In other words, for the repeated use of the skill or repeated performance of the act with more declarative knowledge to facilitate proceduralization, it will be necessary for speakers to pay explicit attention to declarative knowledge stored in their declarative module before task performance and then pay attention to the knowledge again and use it during the ensuing speaking task. Consequently, there should be a greater likelihood of proceduralization occurring. In other words, it is necessary to examine whether helping learners pay attention to, and notice, forms during task repetition may facilitate the occurrence of proceduralization more than simply repeating a task with no opportunity for noticing.

(3) The findings in De Jong and Perfetti are not comparable with those in other studies because the task type and task design are different.

Two types of task have mainly been used in the study of task repetition: a narrative task (Ahmadian & Tavakoli, 2011; Bygate, 1996[17], 2001; Gass et al. 1999) and the poster carousell (Lynch & Maclean, 2001). In addition, the performance of the same task was the main focus of these studies. However, De Jong and Perfetti used a monologue task for learners to talk about a topic and focused on the performance of a new task. In order to compare De Jong and Perfetti (2011) with the results of other studies, the same task type should be used with a similar task design.

The present study thus endeavors to examine the effectiveness of task repetition on proceduralization of linguistic forms by focusing on accuracy and fluency in the same task and a new task. Furthermore, the influence of learners having the opportunity to notice forms during task repetition on proceduralization is examined.

2. The study

2.1 Hypotheses

Three hypotheses were set:

1. When doing the same task as the pretest, greater gains in fluency would be shown for students given the opportunity to notice forms during task repetition than those who were not.
2. When doing the same task as the pretest, greater gains in accuracy would be shown for students

given the opportunity to notice forms during task repetition than those who were not.

3. When doing a new task, different from the pretest, greater gains in fluency and accuracy would be shown for students given the opportunity to notice forms during task repetition than those who were not.

2.2 Procedure

The participants, who voluntarily joined in the study, were 20 English major university students aged 18 to 22. Their TOEIC scores ranged from 550 to 710. They were randomly assigned into two groups (group 1: $N = 10$; group 2: $N = 10$).

Table 1 shows the procedure of this study. Before the pretest was given, three points had been announced orally. The first point was that the purpose of the study was to examine the speaking ability in English of university students, but their names would not be made public. The second point was that the results of their speaking would not be related to their grades in regular classes at all. The third point was that they would have five sessions in total in four weeks. All members of group 1 and half of the members in group 2 followed the procedure outside of the class; the other half of the members in group 2 did so during their English conversation class. Group 1, half of group 2, and the other half of group 2 followed the procedure on different days, respectively (i.e., they were never together).

Both groups had three tests and four training sessions in four weeks. A pretest was given at the beginning of the training sessions, and two posttests were given one week after the training sessions. In each test and training session, they performed a narrative task with a six-strip cartoon extracted from Heaton (1975)[18] (see Appendix 1), which required them to narrate a story using the pictures in order. Posttest 1 was a new task, different from the pretest (a different picture story); posttest 2 was the same task as the pretest (the same picture story). This procedure was modified from Bygate (2001). In his study, after doing a task at the first week, the participants repeated tasks four times in nine weeks. In the 10th week, they performed the same task as in the first week plus a new task different from the first week.

In each training session, each participant was first given an IC recorder and a sheet with a six-strip cartoon. The participants had a quick look at the pictures and made sure that they understood the meaning of each strip. They then narrated a story into the IC recorder for 90 seconds. Soon after telling the story, group 1 simply told the same story again (i.e., no opportunity to notice forms). On the other hand, after telling the story the first time, group 2 was given the opportunity to notice forms through the following three steps. First, they listened back to their story and transcribed it on the back of the sheet with the pictures. While, and/or after, transcribing, they were able to make the best use of self-monitoring and time to retrieve their explicit knowledge from their declarative module. They also looked up linguistic items in their dictionaries. Consequently, they could notice forms which were erroneous and new to them. Second, when self-correction was completed, the participants' transcriptions were submitted, and immediately checked by a native speaker along with the researcher. All errors, except misspellings, in the transcriptions were underlined but left explicitly uncorrected for the participants to notice the errors themselves and returned to each participant. Then the participants had another chance to notice forms which they had not noticed by themselves. Third, after correcting again, each participant submitted the transcription. When a participant still needed correction, explicit and direct feedback using metalinguistic explanation in Japanese was given by the researcher. When all the participants had corrected and understood the errors, they were told to read their story silently three times. They then

Table 1. *Schedule of test and training sessions*

		training sessions									
	pretest	1	[1 week]	2	[1 week]	3	[1 week]	4	[1 week]	posttest 1	posttest 2
group 1	a/a			b/b		c/c		d/d		e	a
group 2	a/(noticing)/a			b/(noticing)/b		c/(noticing)/c		d/(noticing)/d		e	a

Note. The letters refer to the pictures given.

turned over the side of the paper on which they had written their story to look once more at the pictures and told the story again into the IC recorders. After telling the story two times, both groups returned the sheets and IC recorders. [2]

2.3 Analysis

After posttest 2, the sheets used in the tests and the training sessions were given to all participants. Group 1 transcribed their second performance of each task on the back of each sheet while listening back to what they had said on their IC recorders. Group 2 transcribed the first and second performances of all tasks. All transcriptions made by the participants were then compared with those made by the researcher, in order to make sure that all information was transcribed.

After all the transcriptions were made, PRAAT 5.3.09 (Boersma & Weenink, 2012[(19)]) was used to find pauses. The beginning and end of each speech segment was determined first by using the PRAAT function 'To textgrid (silences).' All pause boundaries were checked and adjusted by the researcher as necessary, by listening to the recording and visually inspecting the spectrogram and wave-form. Nonverbal fillers, such as "uh," "ah," "um," and "mmm", were not transcribed. When the filler or silence was 0.20s or longer, it was treated as a pause. This cutoff point was the same as that used by De Jong and Perfetti (2011) although the cutoff point of 0.20s was slightly lower than the 0.25-0.40s that other researchers have used (e.g., Segalowitz & Freed, 2004[(20)]; Towell et al., 1996). In each speech of each participant, the upper limit for pauses was set to 2.5 standard deviations above the mean. A pause longer than the upper limit was replaced by the mean plus 2.5 standard deviations, as in De Jong and Perfetti (2011). Syllables were also counted. Although De Jong and Perfetti (2011) counted false starts as syllables, in this study only words uttered in their entirety were counted as syllables. Words repeated were also counted. To obtain a reliability measure, the number of pauses, the length of each pause, and the number of syllables were all reexamined by two research assistants.

For the measurement of fluency, based on De Jong and Perfetti (2011), the mean pause length, the phonation/time ratio, and the mean length of fluent runs were then calculated as follows:

(*the mean pause length*)
dividing the total length of pauses by the number of pauses;

(*the phonation/time ratio*)
dividing the total time filled with speech (excluding the total length of pauses) by the total time spent speaking (including the total length of pauses) and then multiplying by 100;

(*the mean length of fluent runs*)
dividing the number of syllables by the number of pauses.

On the other hand, accuracy was measured by the ratio of erroneous uses of the target forms as follows:

(*the ratio of erroneous uses of the target forms*)
dividing the number of errors of the target forms by the number of the target forms used and then multiplying by 100.

The frequency of a target form used has a big impact on the ratio of erroneous uses of the form. For example, when no target form is used, the ratio of erroneous uses cannot be measured. When only one target form is used, the ratio becomes 100% if the form is wrong, or 0% if the form is correct. In order to avoid cases whereby a target form is always used less frequently or a form is used inconsistently in each task, articles and verbs were set as the target forms in the present study. It is because articles and verbs are expected always to be used more frequently than other forms and used consistently in each task. Verb errors included tense, word choice, and subject-verb agreement. Article errors covered all types of article uses.

In order to analyze variables for accuracy and fluency, a two-way repeated measures ANOVA and η_{G^2} (generalized eta squared)[3] proposed by Olejnik and Algina (2003)[(21)] to measure effect size were used. For analyzing main effect or simple main effect, η_{G^2} was also used[4]. For analyzing the multiple comparisons, r (for independent t-test) and Δ (for dependent t-test) [5] were used. Power analyses were also conducted by using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009[(22)]).

3. Results

Table 2 shows the descriptive statistics for the mean test scores for the groups. Table 3 shows the

Table 2. *Means and standard deviations of the measures of proceduralization in the tests*

	pretest	posttest 1 (new task)	posttest 2 (same task)
mean number of target forms used			
group 1[a]	27.70 (5.07)	25.40 (4.27)	34.50 (5.75)
group 2[b]	24.40 (6.13)	24.50 (6.91)	24.70 (5.45)
ratio of erroneous uses of target forms			
group 1	40.90 (8.86)	40.80 (11.67)	34.30 (12.60)
group 2	36.10 (12.31)	26.30 (14.52)	16.90 (9.67)
mean length of pauses (in seconds)			
group 1	1.14 (0.29)	1.13 (0.23)	0.90 (0.16)
group 2	0.99 (0.14)	1.23 (0.28)	0.94 (0.21)
phonation/time ratio			
group 1	46.40 (8.42)	43.60 (6.67)	53.00 (7.71)
group 2	46.50 (5.29)	43.60 (5.31)	53.40 (6.78)
mean length of fluent runs (in syllables)			
group 1	2.75 (0.48)	2.23 (0.23)	3.31 (0.58)
group 2	2.47 (0.35)	2.41 (0.49)	3.35 (0.88)

Notes. Values enclosed in parentheses represent standard deviations. [a]$n = 10$. [b]$n = 10$.

Table 3. *Summary of the data analysis of accuracy and fluency in the posttests*

accuracy			
	measure	between comparison	within comparison
posttest 1	number of target forms used	group 1 = group 2	= pretest
	ratio of erroneous uses of target forms	group 1 > group 2	= pretest
posttest 2	number of target forms used	group 1 > group 2	(group 1) > pretest; > posttest 1 (group 2) = pretest; = posttest 1
	ratio of erroneous uses of target forms	group 1 > group 2	< pretest; = posttest 1

fluency			
	measure	between comparison	within comparison
posttest 1	length of pauses	group 1 = group 2	(group 1) = pretest (group 2) > pretest
	phonation/time ratio	group 1 = group 2	= pretest
	length of fluent runs	group 1 = group 2	= pretest
posttest 2	length of pauses	group 1 = group 2	(group 1) = pretest; < posttest 1 (group 2) = pretest; < posttest 1
	phonation/time ratio	group 1 = group 2	> pretest; > posttest 1
	length of fluent runs	group 1 = group 2	> pretest; > posttest 1

summary of the data analysis in the posttests.

3.1 Accuracy

At first, it was examined how much the target forms, articles and verbs, were used. The interaction between group and test was significant with a small sized effect (F (2, 36) = 7.709, p = .002, η_{G^2} = .11, 1–β = 0.99). Looking at the simple main effects by group, significant differences were found only in posttest 2 (the same task) with a large sized effect (t (18) = –3.906, p = .001, r = .68, 1–β = 0.95), and group 1 used more articles and verbs than group 2. In other words, both groups used articles and verbs differently in the same task whereas they used them similarly in the pretest and posttest 1 (a new task). On the other hand, looking at the simple main effects by test, a significant difference was found in group 1 with a large sized effect (F (2, 18) = 16.594, p = .000, η_{G^2} = .39, 1–β = 0.99). Multiple comparisons with a Bonferroni correction (α = .016)

indicated that group 1 used more articles and verbs in the same task than in both the pretest with a large sized effect (t (9) = –4.133, p = .003, Δ = 1.34, 1–β = 0.84) and the new task with a large sized effect (t (9) = –5.386, p = .000, Δ = 2.13, 1–β = 0.97). In other words, group 1 used a different number of articles and verbs in the pretest, the same task, and the new task whereas group 2 used a similar number of the forms at all tests.

Next, the ratio of erroneous uses of the target forms, articles and verbs, was examined. The ratio between the groups was similar in the pretest (t (18) = –1.000, p = .330, r = .23, 1–β = 0.15). However, significant main effect by group was found with a medium sized effect (F (1, 18) = 10.521, p = .005, η_{G^2} = .23, 1–β = 0.74). Group 1 produced a higher ratio with a medium sized effect (t (58) = –3.683, p = .001, r = .44, 1–β = 0.97). In other words, in the pretest, both groups used a similar number of articles and verbs, and the ratios of erroneous uses were also similar between both groups. On the other hand, in the new task, although both groups used a similar number of articles and verbs, the ratio of erroneous uses of those forms by group 1 was higher than by group 2. In the same task, the ratio of erroneous uses of articles and verbs by group 1 was also higher than by group 2, and group 1 used the forms more frequently than group 2.

On the other hand, significant main effect by test was found with a medium sized effect (F (2, 36) = 8.401, p = .001, η_{G^2} = .19, 1–β = 0.99). Multiple comparisons (α = .016) showed that the ratio of erroneous uses of articles and verbs in the same task was lower than in the pretest with a large sized effect (t (19) = 4.492, p = .000, Δ = –1.20, 1–β = 0.95). In other words, although a similar number of articles and verbs were used in the three tests, the ratio of erroneous uses of those forms in the same task was lower than in the pretest.

In summary, the accuracy of use of articles and verbs by both groups improved in the same task, not in the new task; furthermore, group 2 used articles and verbs more accurately than group 1.

3.2 Fluency

In terms of fluency, the data was analyzed based on the aforementioned three measures of proceduralization: mean length of pauses, phonation/time ratio, and mean length of fluent runs.

First, in terms of the mean length of pauses, the interaction between group and test was significant with a small sized effect (F (2, 36) = 3.442, p = .043, η_{G^2} = .05, 1–β = 0.87). Looking at the simple main effects by group, a significant difference was not found. Both groups paused for a similar length of time in the pretest (t (18) = –1.382, p = .184, r = .31, 1–β = 0.25), the new task (t (18) = .829, p = .418, r = .19, 1–β = 0.12) and the same task (t (18) = .453, p = .656, r = .11, 1–β = 0.07). On the other hand, looking at the simple main effects by test, a significant difference was found in group 1 with a medium sized effect (F (2, 18) = 6.733, p = .007, η_{G^2} = .20, 1–β = 0.95) and group 2 with a large sized effect (F (2, 18) = 13.916, p = .000, η_{G^2} = .28, 1–β = 0.95). Multiple comparisons (α = .016) indicated that group 1's pauses were shorter in the same task than the new task with a large sized effect (t (9) = 3.644, p = .005, Δ = –.99, 1–β = 0.73). Group 2's pauses were shorter in the same task than in the new task with a large sized effect (t (9) = –6.104, p = .000, Δ = 1.05, 1–β = 0.99) and also in the pretest than in the new task with a large sized effect (t (9) = –3.626, p = .006, Δ = 1.68, 1–β = 0.73). In other words, in the same task, the length of the pauses of both groups was shorter than in the new task. However, in the new task, the length of the pauses of both groups was not shorter than in the pretest.

Next, in terms of the phonation/time ratio, a significant main effect by test was found with a large sized effect (F (2, 36) = 24.423, p = .001, η_{G^2} = .28, 1–β = 1.00). Multiple comparisons (α = .016) showed that the ratio in the same task was higher than in the pretest with a large sized effect (t (19) = –4.191, p = .000, Δ = .99, 1–β = 0.92) and in the new task with a large sized effect (t (19) = –9.211, p = .000, Δ = 1.64, 1–β = 0.99). In other words, both groups had similar phonation/time ratios in each test. The ratio in the same task was higher than in the pretest and in the new task whereas the ratio in the new task was not higher than in the pretest.

Finally, in terms of the mean length of fluent runs, a significant main effect by test was found with a large sized effect (F (2, 36) = 34.287, p = .000, η_{G^2} = .40, 1–β = 1.00). Multiple comparisons (α = .016) showed that fluent runs in the same task were longer than in the pretest with a large sized effect (t (19) =

–4.873, p = .000, Δ = 1.64, 1–β = 0.98) and in the new task with a large sized effect (t (19) = –8.027, p = .000, Δ = 2.59, 1–β = 0.99). On the other hand, the runs in the new task were similar to in the pretest (t (19) = 2.729, p = .013, Δ = –.67, 1–β = 0.54). In other words, both groups produced similar lengths of fluent runs in each test. The length in the same task was longer than in the pretest and in the new task whereas the length in the new task was not longer than in the pretest.

In summary, the fluency of both groups improved in the same task, not in the new task; however, there was no significant difference in fluency between the groups.

4. Discussion

Hypothesis 1 was that, when doing the same task as the pretest, greater gains in fluency would be shown for students given the opportunity to notice forms during task repetition than those who were not. This hypothesis was not supported. In the same task four weeks after the pretest, both groups performed more fluently than in the pretest. In group 1, who had task repetition, the length of fluent runs became longer, the phonation/time ratio became higher, and the length of pauses became shorter. In group 2, who had not only task repetition but also the opportunity to notice forms, the runs became longer, and the ratio became higher although the length of pauses did not change. However, no significant difference was found between the groups. Fluency was similarly gained whether the opportunity to notice forms was given or not. It is possible to claim that both groups similarly gained fluency because they both had equal measures of task repetition.

This finding that both groups gained fluency contrasts with what Bygate (2001) found. The participants in his study were given the same task 10 weeks after the pretest. They did not improve fluency, though, because the number of unfilled pauses per t-unit was not smaller. Furthermore, the participants in the no repetition group in De Jong and Perfetti (2011) were given a new task 16 to 21 days after the pretest. They also did not improve fluency because the length of pauses, the length of fluent runs, and the phonation/time ratio did not change. On the other hand, the participants in the repetition group in De Jong and Perfetti (2011) did improve fluency because the length of pauses became shorter without negatively impacting the length of fluent runs and the phonation/time ratio. There is one possible reason for such different results. Whereas the participants in Bygate and those in the no repetition group in De Jong and Perfetti were not given task repetition, the participants in the repetition group in De Jong and Perfetti and this study were given task repetition. In other words, task repetition may have led to greater fluency in the same task.

Hypothesis 2 was that, when doing the same task as the pretest, greater gains in accuracy would be shown for students given the opportunity to notice forms during task repetition than those who were not. This hypothesis was supported. In the same task, both groups used the target forms, articles and verbs, more accurately than in the pretest. In group 1, the ratio of erroneous uses of articles and verbs was lower whereas the number of the forms used was larger. In group 2, the ratio of erroneous uses of articles and verbs was lower whereas the number of the forms used did not change. However, the difference between the two groups was significant. Group 2 used articles and verbs more accurately than group 1 although the difference between the groups was not significant in the pretest. A possible reason for this greater accuracy in group 2 than group 1 is that group 2 was given the opportunity to notice erroneous articles and verbs through checking their own transcription and getting feedback from a native speaker and the researcher before repeating the task.

The finding that both groups gained accuracy also contrasts with what Bygate (2001) found. In his study, the participants did not improve accuracy in the same task because the number of errors per t-unit did not decrease. This could be because no task repetition was given in his study. On the other hand, the present study indicates that, although task repetition may lead to greater accuracy in the same task, the opportunity to notice forms may result in even greater accuracy.

Hypothesis 3 was that, when doing a new task, different from the pretest, greater gains in fluency and accuracy would be shown for students given the opportunity to notice forms during task repetition

than those who were not. This hypothesis was partially supported. In a new task, both groups performed less fluently than in the pretest. In group 1, length of fluent runs became shorter, and phonation/time ratio and length of pauses did not change. In group 2, the runs became shorter, the ratio did not change, and the pauses became longer. Furthermore, no significant difference was found between the groups in length of pauses, phonation/time ratio, and length of fluent runs. It can be said that neither group gained fluency. On the other hand, the difference between the two groups was significant in terms of the ratio of erroneous uses of articles and verbs. Although the number of articles and verbs used was similar in both groups, the ratio of erroneous uses of the forms by group 2 was lower than that of group 1. It can be said that the opportunity to notice forms during task repetition may have resulted in this greater accuracy for group 2.

The finding that neither group gained fluency in a new task contrasts with what by De Jong and Perfetti (2011) found, in which the fluency in a new task in the repetition group was greater than in the pretest. There is one possible reason. The language proficiency of the participants was different between their study and this study. This difference is evident in the fluency in the pretest. In their study, the length of fluent runs was 4.50, the phonation/time ratio was 56.94, and the length of pauses was .97. In this study, the length of runs was 2.61, the phonation/ ratio was 46.45, and the length of pauses was 1.065. The participants in their study seem to have been considerably more proficient than those in this study. In other words, whereas the noticing forms during task repetition may have led to greater accuracy in the new task in the present study, fluency may not have improved because the participants were less proficient.

The aggregated results of these three hypotheses may indicate that proceduralization of linguistic knowledge occurred in this study. According to Ellis (2005), evidence for some change in the learner's linguistic knowledge representation can be found in: (1) the learner's use of some previously unused linguistic forms; (2) an increase in the accuracy of some linguistic forms that the learner can already use; (3) the use of some previously used linguistic forms to perform some new linguistic functions or in new linguistic contexts; and (4) an increase in fluency. Each item shows a change in underlying cognitive mechanisms. Such changes in underlying cognitive mechanisms represent proceduralization of linguistic knowledge (De Jong & Perfetti, 2011). Specifically, the increase of accuracy in the same context found in this study is related to items (1) and (2). Furthermore, the increase of accuracy in the new context found in this study is related to (3). Finally, the increase of fluency in the same context found in this study is related to (4). In other words, proceduralization of linguistic knowledge in a speaking task can take place through task repetition, repeated practice of task repetition, and noticing of forms during task repetition.

Three implications can be proposed here:

(1) *Learners should have task repetition in order for proceduralization to occur.*

The long-term effect on the increase in fluency cannot be explained by lexical and structural priming or by planning and attentional resources, and the effect must be attributed to changes in the students' underlying knowledge and processing (De Jong & Perfetti, 2011). Since the same task was conducted four weeks after the pretest in this study, the increase in fluency as well as accuracy may show evidence of proceduralization. Such improvement in both fluency and accuracy, however, conflicts with the trade-off effect between fluency and accuracy proposed by Yuan and Ellis (2003). In their study, the participants either improved fluency not accuracy or improved accuracy not fluency. Such a different finding may have arisen from the fact that the study by Yuan and Ellis (2003) did not have the participants repeat tasks. On a subsequent occasion in task repetition, a speaker's attention to a chunk in declarative module can be freer, and the speaker may be able to fluently or accurately retrieve the chunk. Furthermore, according to ACT-R, through many encounters with the same forms, new production rules are constructed and smaller production rules are collapsed into larger rules, which then become new rules. These new production rules subsequently need to gain strength so as to be able to compete with other, previously existing, production rules. Strength can be gained by repeated practice, so frequent encounters with a form are necessary for

the process of proceduralization. In other words, task repetition is necessary for proceduralization to take place and for not only fluency but also accuracy to improve.

(2) *Learners should be given the opportunity to notice forms during task repetition, in order for accuracy to increase more.*

Proceduralization requires many encounters with the same items. Such frequent encounters will implicitly facilitate learners to make a new production rule. However, such encounters may not always guarantee the making of a new rule, unless learners notice the items. As Schmidt (2001)[23] points out, there is a relationship between noticing a form and frequent processing of the form. When a learner registers a form that is prominent orthographically, phonologically or lexically, the noticed form may be implicitly learned, through being frequently processed later on. Nassaji (2011)[24] also found that repetition of a correct form following recasts had some effect on the accuracy of the form. He then mentioned that repetition of the correct form may indicate that a learner has noticed the form, and the repetition provides the learner with an opportunity to practice the form, which can then enhance its retention and learning. In other words, noticing a form during task repetition and frequent encounters with the form through practice of task repetition are necessary for proceduralization to take place and for accuracy to improve.

(3) *Task type and learner proficiency can influence whether or not, and to what extent, proceduralization will occur.*

De Jong and Perfetti (2011) suggested that task repetition led to repeated use of sentence structures with repeated words, resulting in proceduralization of phrase building and the transfer of fluency improvement to a new task. The participants in this study also had task repetition; however, fluency did not improve in a new task. On the other hand, the participants in this study improved fluency in the same task and produced longer stretches of fluent speech with less pausing for planning. This fluency improvement is greater than that in De Jong and Perfetti because the participants in one of their repetition groups produced longer stretches of fluent speech but with the same length of pauses, while their other repetition group produced only the same lengths of fluent speech albeit with shorter pauses. Both studies used the same measurements of fluency; however, the task types used (a monologue task or a narrative task; the same task or a new task) and the respective participants' proficiencies were different. Such differences may have affected whether or not, and to what extent, proceduralization occurred.

5. Conclusion

The two major findings in this study were:

1. Through task repetition, proceduralization of linguistic knowledge with greater fluency and accuracy will occur in the same task.
2. With an opportunity to notice language forms during task repetition, proceduralization of linguistic knowledge with greater accuracy will occur in the same task and a new task.

"Proceduralization is considered a slow process that requires many encounters with the same items" (De Jong & Perfetti, 2011, p. 562). The many encounters with articles and verbs through task repetition in this study might not have developed procedural knowledge enough for it to be transferred to a new task. However, such encounters might nevertheless have developed procedural knowledge enough for it to be used in the same task. In order that learners develop procedural knowledge little by little and enact the slow process of proceduralization, it is useful to have learners repeat tasks for frequent encounters with the same forms.

However, there are at least three limitations in this study. The first is that the time allocated to notice the target forms and the time used to reflect on the noticed forms before repeating the same task were not necessarily equal to the time given to the participants in group 2. The participants who had spoken less and spent less time transcribing their speaking were able to have more time to notice the erroneous forms through self-monitoring and a native speaker's check than those who had spoken more. Those participants who had spoken less and finished noticing the erroneous forms sooner were also able to have more time to attend to the forms before repeating the task than those who had

spoken more. The difference in time allocated to notice the target forms and used to reflect on the noticed forms before repeating the same task might have influenced the degree of proceduralization or experience of form-meaning connection necessary for acquiring procedural knowledge.

The second limitation is that the forms targeted in the study were not sufficiently focused. Verb errors included tense, word choice, and subject-verb agreement. Article errors covered all types of article uses. Therefore, it was not clarified if a specific feature of a form was used erroneously, the feature was noticed, and/or the form with the same feature was used repeatedly in a new task and in the same task. In the process of proceduralization, it is necessary for learners to experience first the form-meaning connection in an exemplar repeatedly, and later experience the same form-meaning connection in many exemplars repeatedly and then generalize the connection. A specific feature of a focused form should have been targeted in order to examine if task repetition can facilitate proceduralization.

The last limitation is that the nature of task repetition in the two groups was not necessarily the same. In each session, group 1 simply repeated the same task twice. On the other hand, group 2 did the first performance, transcribed the first performance for noticing and then undertook the second performance. In other words, group 2 repeated the same task three times in each session. The different results between the two groups might have been caused by the different frequency of repetition of the same task, not by the opportunity of noticing erroneous forms.

Considering such limitations, further research is necessary to expand the potential of task repetition to facilitate the proceduralization of linguistic knowledge. First, it is necessary to examine whether other types of activities for noticing forms, other than self-monitoring plus feedback, are effective, such as the study of Ahmadian and Tavakoli (2011) combining task repetition with careful on-line planning. Second, it is necessary to examine the effects of various types of task repetition. One is repeating a task after a while as in Bygate (2001), or immediately repeating a task numerous times as in this study, De Jong and Perfetti (2011), and Lynch and Maclean (2001). Another variable is repeating a task after doing different types of task as in Bygate (2001), or after doing the same type but a new task as in this study and De Jong and Perfetti (2011). Third, it is necessary to focus more on the influence of variables, such as task type and learner proficiency, on the proceduralization of linguistic knowledge. As a possible reason for the different results between this study and De Jong and Perfetti (2011) regarding fluency gain in a new task, the lower proficiency of the participants in this study was noted. It should be examined if there is a threshold level to facilitate proceduralization of linguistic knowledge with fluency in a new task through task repetition. Fourth, it is necessary to examine the long-term effect of the combination of task repetition and noticing by a delayed posttest. De Jong and Perfetti (2011) gave a new task as an immediate posttest one week after the last training session and another new task as a delayed posttest three weeks after the last training, so giving a delayed posttest might clarify whether the effectiveness was temporary or will be retained.

Notes

1. The procedure of the poster carousel is as follows:
(1) Participants are paired up and each pair is given a different research article. Then they make a poster based on the article.
(2) The posters are displayed around a room. From each pair (A and B), one participant (A), the host, stands beside his or her poster, waiting to receive visitors asking questions. The B participants visit the posters one by one, clockwise. They ask questions about each poster. The host responds to questions. A limited time is allowed for discussing each poster.
(3) When the B participants arrive back at base, they stay by their poster, and then the A participants go visiting.
(4) Once the second round is completed, there is plenary discussion of the merits of the posters and the teachers provide feedback on general language points.

2. After posttest 2, two points had been announced orally. The first point was that the data which had been recorded in each IC recorder might be used later in presentation at conferences or in papers, but their names would not be made public. The second point was that those who want their data to be omitted should declare so now or later. The participants who did not declare their desire to

have their data omitted were then judged to have approved of the use of their data.

3. η_P^2 (partial eta squared) may become problematic when there are one or more repeated measures factors because it cannot be directly compared across studies with between- and within-subjects designs (Bakeman, 2005[25]). η_G^2 was calculated by $SS_A / (SS_T - SS_B - SS_{AB})$ for factor A, $SS_B / (SS_T - SS_A - SS_{AB})$ for factor B, and $SS_{AB} / (SS_T - SS_A - SS_B)$ for factor AB. The interpretations of η_G^2 were based on Bakeman (2005): $|.02| \leq$small$< |.13|$; $|.13| \leq$medium$< |.26|$; $|.26| \leq$large.
4. Referring to Olejnik and Algina (2003), $\eta_G^2 = SS_A / (SS_A + SS_S + SS_{S/A})$
5. $r = \sqrt{t^2 / (t^2 + df)}$ $\Delta = |\text{Mean1}| \text{Mean 2}| / |\text{SD1}|$ The interpretations were based on Cohen (1988)[26]: r ($|.10| \leq$small$< |.30|$; $|.30| \leq$medium$< |.50|$; $|.50| \leq$large) and Δ ($|.20| \leq$small$< |.50|$; $|.50| \leq$medium$< |.80|$; $|.80| \leq$large).

References

(1) Lynch, T., & Maclean, J. (2001). Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 141–162). Harlow: Longman.

(2) Ahmadian, M., & Tavakoli, M. (2011). 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.

(3) 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: Longman.

(4) Gass, S., Mackey, A., Fernandez, M., & Alvarez-Torres, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*, 549–580.

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

(6) Ellis, R. (2005). *Planning and task performance in a second language*. Amsterdam: John Benjamins Publishing Company.

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

(8) Anderson, R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

(9) Anderson, R., Bothell, D., Byrne, D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036–1060.

(10) Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*, 145–164.

(11) Towell, R., Hawkins, R, & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 84–119.

(12) Nation, P. (1989). Improving speaking fluency. *System, 17*, 377–384.

(13) Kim,Y., & McDonough, K. (2008). Learners production of passives during syntactic priming activities. *Applied Linguistics, 29*, 149–154.

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

(15) 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: John Benjamins Publishing Company.

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

(17) Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis, & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Oxford: Heinemann.

(18) Heaton, J. (1975). *Beginning composition through pictures*. Harlow: Longman.

(19) Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer. http://www.fon.hum. uva.nl/praat/

(20) Segalowitz, N., & Freed, F. (2004). Context, contact, and cognition in oral fluency acquisition. *Studies in Second Language Acquisition, 26*, 173–

199.
(21) Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*, 434–447.
(22) Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
(23) Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language acquisition* (pp. 3–33). Cambridge: Cambridge University Press.
(24) Nassaji, H. (2011). Immediate learner repair and its relationship with learning targeted forms in dyadic interaction. *System, 39*, 17–29.
(25) Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*, 379–384.
(26) Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Appendix 1. A six-strip cartoon used in the present study