THE JOURNAL OF ASIA TEFL
Vol. 18, No. 4, Winter 2021, 1108-1127
http://dx.doi.org/10.18823/asiatefl.2021.18.4.3.1108

The Journal of Asia TEFL

http://journal.asiatefl.org/

e-ISSN 2466-1511

Revised 4/3/2 Task: Fluency Training with Formulaic Language in the EFL Classroom

Chie Ogawa
Kyoto Sangyo University, Japan

The purpose of this study was to examine to what extent additional interventions to the 3/2/1 task help EFL learners develop speaking fluency. Participants were 48 university students in Japan. The 3/2/1 task was implemented 11 times in one 15-week academic semester. The participants were divided into three groups including 1) comparison group, 2) input enhancement group, and 3) input + peer-check group. Speaking data were recorded in weeks 2 and 14, transcribed, and analyzed based on five fluency measures. The pre-/post-recording data showed significant gains in mean length of runs in both experimental groups and in phonation time ratio in the input + peer-check group. Further analyses indicated that the increase in the mean length of runs is associated with the frequency of target formulaic language usage. The implications of fluency promoting instructions in EFL classrooms are discussed.

Keywords: instructed SLA, speaking fluency, task-based language learning, formulaic language, repetition

Introduction

There has been a growing interest in fluency development practice in second language acquisition. In an English as a foreign language (EFL) context such as Japan, wherein second language (L2) learners have limited access or opportunities to use the target language outside the classroom (Leeming & Harris, 2020), pedagogical intervention for fluency development inside the classroom is essential to help learners become better communicators according to *Four Strands* (Nation, 2007). In most language classrooms, however, language teachers tend to focus more on speaking practice in general than on fluency-specific features, such as pauses, repairs, and speed (Tavakoli & Hunter, 2018). Indeed, pedagogic activities designed to promote a narrow sense of fluency (Lennon, 1990; Segalowitz, 2010; Tavakoli & Skehan, 2005) are often neglected in language teaching (Rossiter et al., 2010; Tavakoli & Hunter, 2018). Therefore, this study examined the effectiveness of pedagogical intervention to enhance L2 learners' oral fluency in the EFL context.

There are several ways to enhance speaking fluency (Foster, 2020; Tavakoli & Hunter, 2018). One well-known repetition task is the 4/3/2 task (Boers, 2014; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016; Tran & Saito, 2021; Wood, 2009). To date, the 4/3/2 task has been widely used in many EFL/ESL classrooms. The task is believed to help L2 speakers express their ideas more efficiently, quickly and smoothly due to the repetition under shrinking time. Another effective way to enhance speaking fluency is teaching formulaic language (Boers et al., 2006; Tavakoli & Uchihara, 2020; Wood, 2009, 2010, 2015). Proficiency in formulaic language can help learners to speak fluently because it can be retrieved faster than sentences that are generated verbatim under real-time conditions (Boers et al., 2006;

Tavakoli & Uchihara, 2020; Wood, 2009). The current study explores the extent to which utilizing the 4/3/2 task and teaching formulaic language can enhance L2 speakers' speaking fluency.

Literature Review

4/3/2 Task for Fluency Development

In the 4/3/2 task, L2 learners tell a monologue about the same topic for four minutes, then three, and finally two. When the L2 speakers tell the same topic three times with increasing time pressure to perform more quickly, they must speak faster. The task has three important features: repetition, a reduction in time, and a change of audience (Nation, 1989). These features directly affect fluency by encouraging L2 speakers to focus on the meaning under a time constraint.

The effects of the 4/3/2 tasks have been empirically investigated by several researchers. De Jong and Perfetti (2011) investigated whether repeating the same topic in the 4/3/2 task would lead to a long-term increase in oral fluency with 24 adult English as second language (ESL) learners in the United States. The participants were randomly assigned into repetition, no-repetition, and control groups. The repetition group had the same topics, whereas the no-repetition group spoke on three different topics in the 4/3/2 task. The participants' oral data for 2-minute personal-story monologue on different topics were collected at pretest, posttest, and delayed posttest. Their findings showed that repeating the same topics during the 4/3/2 tasks is more effective than talking on different topics for developing oral fluency and retaining the effects of fluency over four weeks. Fluency improvements during the 4/3/2 task were the result of proceduralization by the participants who repeated the same topic (De Jong & Perfetti, 2011).

To investigate whether a shrinking time condition can promote syntactic accuracy, Boers (2014) tested learners' performance under a time-constant and a time-shrinking condition. Ten adult ESL learners were asked to select two topics they felt comfortable talking about. To counterbalance the task order, five participants did the 4/3/2 task first, while the other five did the 3/3/3 activity first. Mean quantitative changes in complexity, accuracy, fluency (CAF) indices between the first and third deliveries were summed for the two task conditions. Boers (2014) compared mean changes on CAF indices between the first and third deliveries, with results showing that learners in the shrinking time condition improved their fluency more significantly. The similar finding was reported in a case study with a L2 learner in Singapore (Bui, 2021).

Thai and Boers (2016) conducted a similar study wherein they examined the 4/3/2 speaking task with and without time pressure. The participants were 20 tenth grade EFL students in Vietnam who talked about the same topic—their favorite movie. Ten participants were in the 3/2/1 condition, while the other ten participants were in the 2/2/2 condition. The researchers analyzed all 60 speeches (20 participants x 3 deliveries) using CAF indices. The results indicated that oral fluency (syllables per minute) improved significantly under the time-shrinking condition (3/2/1); however, there were no significant changes under the 2/2/2 condition.

Previous studies (Boers, 2014; Bui, 2021; De Jong & Perfetti, 2011; Thai & Boers, 2016) did not investigate additional remedial activity during the 4/3/2 task. However, Tran and Saito (2021) examined the effectiveness of accuracy enhancement (AE) during the 4/3/2 task. Thirty-six Vietnamese EFL learners were divided into three groups (3/3/3, 4/3/2, 4/3/2+AE). Participants in the 4/3/2+AE group received corrective feedback on the accuracy of past tense. In line with previous studies (Boers, 2014; Thai & Boers, 2016), the finding showed that engaging in the 4/3/2 task helped participants become more fluent, but their accuracy did not improve. In contrast, participants in the 4/3/2+AE group improved both fluency and accuracy over three sessions of the 4/3/2 task. This study is noteworthy because it was one of the first to examine additional remedial activity in the 4/3/2 task.

In sum, studies mentioned previously provide the following evidence. First, between the first and third delivery, the 4/3/2 task helped L2 learners to immediately improve oral fluency because repetition of the

same topic and time pressure promote automatization and proceduralization of knowledge (Boers, 2014; Thai & Boers, 2016). Second, the 4/3/2 task helped maintain fluency development and further enhanced other topics' transferability over time (De Jong & Perfetti. 2011; Tran & Saito, 2021). Third, additional intervention (corrective feedback) during the 4/3/2 task helped learners improve accuracy and fluency for both immediate changes between the first and third deliveries and in the longer span after the three sessions of the 4/3/2 tasks (Tran & Saito, 2021). Table 1 summarizes previous studies on the 4/3/2 task.

TABLE 1
Summary of Previous Studies of the 4/3/2 Task

Studies	N	Participants' proficiency level	Setting	Procedures	Findings about fluency
De Jong and Perfetti (2011)	24	ESL adult learners in USA High intermediate, (60–79 Michigan Test of English Language Proficiency)	In regular ESL class hours, USA	Three sessions of the 4/3/2 treatment pretest - posttest- delayed posttest Repeating the same topic vs. talking about different topic each time	The repetition group improved fluency over time Fluency gains were transferred to new topics in delayed posttest (after three weeks)
Boers (2014)	10	ESL adult learners in NZ Intermediate to advanced (self-reported IELTS scores ranged from 5 to 7.5)	Laboratory conditions with other participants in pairs	One single session with counter-balanced design shrinking condition (4/3/2) vs. constant condition (3/3/3)	Immediate fluency improvement between 1st and 3rd delivery. 4/3/2 > 3/3/3
Thai and Boers (2016)	20	10th grade high school learners in Vietnam (no proficiency level reported)	In regular EFL classes	One single session of the treatment shrinking condition (3/2/1) vs. constant condition (2/2/2)	Immediate fluency improvement between 1st and 3rd delivery. 3/2/1 > 2/2/2
Tran and Saito (2021)	36	EFL learners in university in Vietnam: pre-intermediate (TOEIC 360-690)	Individual session with researcher in a laboratory	Three sessions of the treatment shrinking condition (4/3/2) vs. constant condition (3/3/3) vs. Intervention condition (4/3/2) + AE (accuracy enhancement)	Immediate fluency improvement between 1st and 3rd delivery (4/3/2). long-term fluency improvement was seen in 4/3/2 and 4/3/2 + AE after the 3 sessions.

From these previous findings, we can distinguish two types of fluency improvement in 4/3/2 studies. Researchers examined fluency improvement in different time spans, that is, immediate changes between first and third deliveries in a single session of the 4/3/2 task (Boers, 2014; Bui, 2021; Thai & Boers, 2016) or longer-term fluency improvement throughout repetitive practices of the 4/3/2 task (e.g., after three sessions of 4/3/2 training in DeJong & Perfetti, 2011; Tran & Saito, 2021). From the practitioners' perspective, the latter would reveal more helpful information because language teachers usually evaluate L2 speakers' improvement over a semester rather than over a one-day session. Additionally, some studies were conducted in laboratory settings. Previous researchers have already acknowledged this issue (Boers, 2014; Tran & Saito, 2021); thus, a longer-term design for tracking L2 learners' speaking fluency over multiple 4/3/2 sessions in classroom settings is needed.

Another gap in the literature is that few studies have examined additional intervention during the 4/3/2 task. For example, as De Jong and Perfetti (2011) recognized, it is uncertain which linguistic forms are proceduralized through the 4/3/2 task because no intervention has been provided before or during the task. They suggested providing learners with model input and encouragement to use the input in their own speech, which might result in greater lexical use because in L2, learners are encouraged to mine from input for exemplars to integrate into their own speech (Boers, 2014; Thai & Boers, 2016). Although Tran and Saito (2021) examined the effectiveness of corrective feedback on enhancing L2 learners' use of correct past tense, little is known about the effectiveness of other awareness-raising activities during the 4/3/2 task.

The Journal of Asia TEFL
Vol. 18, No. 4, Winter 2021, 1108-1127

Formulaic Language Instruction and Oral Fluency

According to Wray (2002), formulaic language refers to the formulaic sequences or multiword units that are or appear to be prefabricated. The examples include collocations (*take advantage*), idioms (*take away your breath*), and compounds (*bullet point*). These units are stored and retrieved whole from memory during use, rather than being subject to grammatical analysis (Wray, 2002). If learners are able to process formulaic language automatically, they can use more attentional resources for tasks such as retrieving lexical items or creating a syntactically accurate utterance (Segalowitz, 2003).

Empirical studies have demonstrated a relationship between formulaic language usage and oral fluency (Boers et al., 2006; Tavakoli & Uchihara, 2020). For example, greater use of high frequency formulaic language is related to speed fluency (Tavakoli & Uchihara, 2020). On the other hand, limited use of formulaic languages was due to L2 learners' lack of proficiency (Zipagan & Lee, 2018). Foster (2020) highlights the need to explore the extent to which it is feasible to work on developing oral fluency and formulaic language among classroom learners.

Boers and his colleague (2006) employed a pedagogical intervention that emphasized noticing formulaic language in a text. In their study, Belgian EFL learners worked in pairs to underline or highlight useful phrases in a text over 22 teaching hours. The result showed that participants in the awareness-raising group were perceived to be more fluent speakers than the control group. Their findings suggest that noticing input helped participants build a repertoire of formulaic language that contributed to their oral fluency's improvement. However, as the authors acknowledged, productive practice in real-time communication is needed. Gatbonton and Segalowiz (1988) suggested a productive practice to incorporate repetition and rehearsal of formulaic language into a communicative task. Unlike the tradition of repeating a target form in monotonous drills, they called this process *creative automatization* for L2 learners to create appropriate utterances based on the communicative situation.

Some researchers have examined using formulaic language instruction in the input and output stages (Wood, 2009; Thomson, 2017). Wood (2009) conducted a case study of a formulaic language workshop with a female Japanese ESL student in Canada. Sessions included (a) an input stage, (b) an automatization stage, (c) a practice and production stage, and (d) a free talk stage. In the input stage, the participant listened to native English speakers' personal stories. The instructor drew attention to formulaic language and commented on its linguistic and discourse functions. In the automatization stage, the participant shadowed the recorded model at least eight times. The participant then performed a dictogloss activity by listening to the sentences with the formulaic language taken from the input passage. In the practice and production stage, the participant did the 4/3/2 task wherein she told personal narratives. In the free talk stage, she talked about assigned topics. The participant gained 26.30% in mean length of run (MLR) between pretests and posttests after a six-week training session. Thomson (2017) replicated Wood's (2009) fluency workshop with 73 EFL learners at a Japanese university. The result showed that the fluency workshop helped learners increase fluency over six weeks. These studies (Wood, 2009; Thomson, 2017) support the idea that instructional approaches that provide learners with ample output practice to foster the proceduralization of formulaic language possibly hold greater promise for oral fluency development than 4/3/2 speaking practice alone.

Nevertheless, more studies are needed to examine in real classroom settings because little has been found regarding effective ways to utilize the 4/3/2 task with target formulaic language (TFL) in the EFL classroom. This study proposes a classroom-based study that employs additional awareness-raising activities (input enhancement and peer-check) during the 4/3/2 task. Specifically, this study is guided by the following research questions:

1) To what extent do the EFL learners in a Japanese university develop their oral fluency through the 3/2/1 task with a pedagogical intervention to use the target formulaic language?
2) Is there a relationship between usage of the target formulaic language and fluency development?

Methods

Classroom Settings

This study was conducted in an English discussion program at a private Japanese university. All first-year students were required to take a weekly 90-minute English discussion course in their first year of study for 14 weeks in the spring semester (April–July), from which the data were collected. The author was the teacher of the English discussion class. The average class size was small (7–9 students), and the classes were held only in English. The main objective of the course was for students to learn how to effectively participate in group discussions and develop speaking fluency using the TFL.

Participants

The participants were 48 first-year Japanese university students; 18 male and 30 female participants. The learners' TOEIC scores are between TOEIC 350 and 550 (see Table 2). The three subgroups (comparison, input enhancement, input + peer-check group) are explained in detail in the section below.

TABLE 2
The Participants' Information

	Comparison group (n = 12)	*Input enhancement group (n = 13)*	*Input + Peer group (n = 21)*
Mean age	18.2 years	18 years	18.2 years
Gender	Male = 3	Male = 7	Male = 8
	Female = 9	Female = 7	Female = 13
TOEIC score	*M* = 519.17	*M* = 479.29	*M* = 480.00
	SD = 30.66	*SD* = 38.17	*SD* = 57.20

Target Formulaic Language

The study aimed to raise awareness of TFL rather than teach unknown phrases. The 10 types of formulaic language were chosen from textbook instruction. Each sequence comprised two to six words. TFL included (a) stating opinions (e.g., *In my opinion, Personally speaking I think, I'm not sure but I think*); (b) giving reasons (e.g., *It is mainly because, One reason is, Another reason is*…); and (c) giving examples (*For example, For instance, One example is, Another example is*). All participants studied the same TFL for the course's group discussion (Appendix A). Target phrases were used as they were shown in the original textbook that the university had created to facilitate 16-minute group discussions. These target phrases were derived from Dörnyei and Thurrell's (1994) direct approach to conversation instruction, and Kehe and Kehe's (1994) speaking text *Conversation Strategies*. These TFL types were chosen because to linguistic form (TLF) should be multi-situational, that is, usable for many topics with little or no modification in an opinion-based task, followed with the rationale by Gatbonton and Segalowiz (1988). Another reason is that these TFLs are expected to help participants lengthen their speaking time; this course's first-year university students have generally expressed difficulty in presenting a three-minute monologue.

Procedure of the 3/2/1 Task

The 3/2/1 task was implemented every week in class. Instead of 4/3/2, a shorter 3/2/1 task was used due to the participants' lower proficiency level (e.g., Thai & Boers, 2016). At the beginning of every class, participants engaged in the 3/2/1 task. The participants formed pairs of speakers and listeners. One speaker discussed a particular topic for three minutes, retold the information the second time in two minutes, and then repeated it in one minute. The listeners did not interrupt the speakers by making

comments or asking questions. The speakers completed a full 3/2/1 task with three different listening partners each round before switching roles.

The participants' oral performance during the 3/2/1 task was analyzed twice, once in week 2 (pretest) and once in week 14 (posttest) during an academic semester to understand participants' oral fluency development throughout the academic semester. The participants recorded their 3/2/1 task production individually using an IC recorder on the recording days in the same classroom during the English discussion class. The recording usually took 15 minutes to complete. The questions provided for the task were different to avoid repetition effect as shown in Table 3.

TABLE 3
3/2/1 Task Recording Questions

Recording	Questions
Pretest	Club activity Do you think doing a club activity is a good idea for students? Have you ever joined a club before? What did you learn from your experiences? Why did you choose your club in this university?
Posttest	Learning English Do you think learning English is important for students? Do you think studying abroad is a good idea for students? What are other good ways to improve your English skills?

Pedagogical Treatment

The participants were divided into the three groups: 1) comparison, 2) input enhancement, and 3) input + peer-check, with each group comprising two or three intact classes. The class combinations were randomly sorted based on the participants' TOEIC scores so that their English levels were similar for each group. The comparison group started the 3/2/1 task immediately after the teacher showed the today's topic.

The input enhancement group received the teacher-led model passage (Appendix B) using the formulaic language with the handout. The teacher-modeled passage was displayed on the handout with the TFL, which was enhanced with underlining. All types of TFL (e.g., opinion, reason, and example) were provided in each lesson's model input. Although participants did not study some TFL until later in the semester (e.g., participants were introduced to the example function in Week 6), they were exposed to the teacher-modeled passage earlier during the 3/2/1 task. While the teacher read the passage aloud, the participants read it silently, taking less than one minute. After the participants finished reading, they were given two minutes to generate ideas by writing phrases in English on the handout. Participants had not been allowed to look at the handout when they performed the 3/2/1 task. The participants were encouraged to organize their speech by supporting their opinions with reasons and examples, so they would be able to speak longer. The total length of the treatment was three minutes (one minute of reading the teacher-modeled passage and two minutes of pre-task planning).

The input + peer-check group received the same treatment as the input enhancement group and the following additional pedagogic intervention. While the speakers engaged in the 3/2/1 task, the listeners exerted pressure to use the target forms. This intervention was designed to raise awareness and encourage the speakers to use the TFL. Listeners checked to see if the speakers used the TFL during the 3/2/1 task and gave feedback on the usage (Appendix C).

Table 4 demonstrates the formulaic language-training schedule. For every lesson, participants engaged in the 3/2/1 task related to the lesson topic. For example, the 3/2/1 task questions on week 5 were: *Do you think going to university is important? Why did you decide to come to university? What are your future plans after you graduate from university?*

TABLE 4
The Target Formulaic Language-training Schedule

Week	*Lesson Topic*	*Formulaic Language Taught*
1	Practice and introduction	
2	Pretest	Opinion (*In my opinion, Personally speaking I think, I'm not sure but I think*)
3	Communication	Reason (*It's mainly because, One reason is, Another reason is*)
4	Education	
5	Education	
6	Environment	Example (*For example, For instance, One example is, Another example is*)
7	Environment	
8	Social issues	
9	Social issues	
10	Technology	
11	Technology	
12	Values	
13	Values	
14	Posttest	

Fluency Measures

Only the two-minute performance of the recorded 3/2/1 speaking data was analyzed because it was considered as an appropriate length (De Jong & Perfetti, 2011). The speech data were transcribed in the following manner. First, the recorded data, including fillers and self-repetitions, were transcribed. Next, transcriptions were double-checked by a research assistant. Discrepancies were discussed until we reached an agreement. Speech data comprising of 192 minutes were transcribed (2 minutes × [48 participants × 2 times]).

Utterance fluency can be categorized into three groups: breakdown, repair, and speed fluencies (Tavakoli & Skehan, 2005). This study used five measures of fluency: 1) mean length of pauses (MLP), 2) number of repairs and repetition, 3) mean duration of syllables (MDS), 4) MLR, and 5) phonation time ratio (PTR). MLP is related to breakdown fluency, number of repairs, and repetition for repair fluency, and mean duration, to speed fluency. MLR and PTR are a combination of breakdown and speed fluencies.

Mean length of pauses

Silent and filled pauses are two kinds of pauses. First, to define silent pauses, the threshold was set at 300 ms (De Jong & Bosker, 2013; Thai & Boers, 2016), given that the participants' oral proficiency level is not advanced. Second, nonverbal fillers such as *uh*, *ah*, and *um* were treated as silence (De Jong & Bosker, 2013; De Jong & Perfetti, 2011; De Jong et al., 2015); this is because participants used fillers or kept silent while thinking. The word *pauses* in this study includes silent and filled pauses. Length of pauses was measured using the Praat speech analysis software (Boersma & Weenink, 2009).

Number of repairs and repetition

Repair fluency includes (a) false starts, (b) reformulation, and (c) the repetition of words or phrases (Tavakoli & Skehan, 2005, p. 255). Fillers were not included in the analysis of repair fluency to avoid an overlap with breakdown fluency. False starts are words left as incomplete clauses and followed by a new start involving different lexis and syntax.

Mean duration of syllabus

As a measure separate from other disfluency components such as pauses and repairs (De Jong et al., 2015), speed fluency was calculated as the MDS, which was calculated as speaking time divided by the number of syllables. This measure was used in previous studies to avoid confounded measures (Bosker et al., 2013; De Jong & Bosker, 2013; De Jong et al., 2015). When analyzing the MDS, speaking time is used after excluding pauses.

Mean length of runs

MLR was calculated as the mean number of syllables produced in an utterance between pauses (total number of syllables divided by number of runs). A run is a fluent sequence between two silent pauses, which were identified using a cut-off rate of 300 ms. The number of runs was calculated by adding 1 to the number of pauses. For example, if there were seven pauses, then there were eight runs: 7 +1 = 8. Then, the total number of syllables was divided by eight.

Phonation time ratio

The PTR was calculated as the total length of phonation time (time spent speaking) divided by the total response time a participant spent speaking (two minutes). The total length and the number of silent pauses was determined using the cut-off rate of 300 ms. Phonation time was determined by subtracting the total time of silent pauses from the total response time (e.g., 120 seconds in total − 30 seconds silent time = 90 seconds). Table 5 shows the calculations for the fluency measures.

TABLE 5
Fluency Measurements

Types	*Specific measures*	*Calculation*
Breakdown fluency	Mean length of pauses [MLP]	Sum of pauses / number of pauses
Repair fluency	Number of repairs, repetitions, and false start	Number of repairs, repetitions and false start / spoken time
Speed fluency	Mean duration of syllables [MDS]	Spoken time / number of syllables
Combination	Mean length of run [MLR]	Number of syllables / number of runs
	Phonation time ratio [PTR]	Spoken time / total time

Note. Spoken time means phonation time spent on speaking without silent pauses and fillers. Spoken time includes repairs.

All the transcriptions were double-checked by a research assistant. Syllable count was calculated using software on a website (http://www.syllablecount.com), so inter-rater reliability was not used.

Analysis

A repeated-measures MANOVA was run to investigate whether there were significant differences in fluency gain between the three groups. The dependent variables were the five fluency measures: 1) MLP, 2) number of repairs and repetition, 3) MDS, 4) MLR, and 5) PTR. The independent variables were Time (pre and posttests) and Group (comparison, input enhancement, input + peer-check).

Separate ANOVAs were then used to compare the participants' fluency changes between pre and posttests. To avoid committing a Type I error, a Bonferroni adjustment was used by dividing the alpha level of.05 by the five dependent variables used (p = .01). Before conducting the repeated-measures ANOVAs, the assumptions were checked for each group two times. No univariate outliers were found, and the assumption of normality was met.

Participants showed different levels of performance initially because of the nature of the intact classroom. One-way ANOVA to compare mean differences between groups for the pretest indicated that two of the five fluency measures revealed significant differences: MLR, $F(2, 45) = 6.81$, $p = .003$, partial $\eta^2 = .23$, and PTR, $F(2, 45) = 5.60$, $p = .007$, partial $\eta^2 = .19$. Post-hoc analysis indicated that the comparison group showed longer MLR and higher PTR than the input + peer-check group before the pedagogical intervention. Therefore, the results should be interpreted with caution.

Results

Table 6 illustrates the descriptive statistics for all the measures of fluency at pretests and posttests. The participants in each group increased the MLP but could produce longer runs and higher PTRs.

TABLE 6
Descriptive Statistics for Fluency Development

	Comparison (n = 12)			Input enhancement (n = 13)			Input + Peer feedback (n = 21)		
	pretest	posttest		pretest	posttest		pretest	posttest	
	M *(SD)*	*M* *(SD)*	gain %	*M* *(SD)*	*M* *(SD)*	gain %	*M* *(SD)*	*M* *(SD)*	gain %
Pauses	0.75 -0.21	1.05 -0.15	40.00%	0.71 -0.18	2.29 -0.38	223%	0.59 -0.13	1.57 -0.43	166.10%
Repairs	12.15 -9.07	11.69 -2.93	-3.79%	10.5 -5.88	10.64 -5.72	1.33%	7.71 -5.03	8.05 -5.01	4.41%
MDS	0.28 -0.03	0.3 (-0.03)	7.14%	0.29 -0.04	0.29 -0.03	0%	0.3 -0.05	0.31 -0.03	3.33%
MLR	4.88 -0.75	5.05 -0.8	3.48%	4.58 -0.72	5.31 -0.86	15.94%	3.95 -0.76	5.14 -0.93	30.13%
PTR	53.97 -8.51	60.83 -4.33	12.71%	52.31 -9.65	55.61 -7.85	6.31%	45.5 -6.11	51.85 -8.88	13.96%

Note. MDS = Mean Duration of Syllable, MLR = Mean Length of Run, PTR = Phonation Time Ratio

The repeated-measures MANOVA analysis showed multivariate significant differences for Time (Wilks' Lambda = .70; $F = 19.32$, $p < .001$; $\eta^2 = .30$), but no significant difference for the interaction between Time and Group (Wilks' Lambda = .97; $F = .59$, $p = .559$; $\eta^2 = .26$). A series of one-way repeated-measures ANOVAs was conducted with the factor being time (two levels: pretest and posttest) and the dependent variables were the five fluency measures. A Bonferroni-adjusted alpha level of .01 (.05/5) was used for the paired *t*-test. Using Plonsky and Oswald's (2014) benchmarks for Cohen's *d*, $d = .60$ is considered as small, $d = 1.00$ as medium, and $d = 1.40$ as a large effect size.

The results of the *t*-tests, comparing fluency changes from pre to posttests, are presented in Table 7. All the groups (comparison, input enhancement, input + peer feedback) increased pause length with large effect size (Cohen's $d = 1.68, 2.14, 3.54$ respectively). None of the groups increased the number of repairs. Two experimental groups significantly increased the MLR. The input enhancement group's effect size was small (Cohen's $d = .79$), and the input + peer-check group's effect size was medium (Cohen's $d = 1.31$). Only input + peer-check group significantly increased PTR (Cohen's $d = .76$).

TABLE 7
Paired Sample T-test Comparing Fluency Development from Pre to Posttests

Fluency measures	*Comparison group (n = 12)*			*Input enhancement (n = 13)*			*Input + Peer feedback (n = 21)*		
	t	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*
Pauses	−4.89	.000*	1.68	−4.48	0.001	2.14	−9.4	.001*	3.54
Repairs	−.21	0.84	—	0.07	0.94	—	0.34	0.74	—
MDS	−3.13	.009*	0.66	0.16	0.88	—	−.27	0.79	—
MLR	0.71	0.49	—	−2.96	.01*	0.79	−5.98	.000*	1.31
PTR	−2.97	0.01	0.9	1.53	0.15	—	−3.33	.003*	0.76

Note. MDS = Mean Duration of Syllable, MLR = Mean Length of Run, PTR = Phonation Time Ratio. *$p < .01$.

Both experimental groups increased MLR with small-to-medium effect size. To understand whether the frequency of the TFL use was related to the increase of MLR, the average usage of the TFL per person was calculated by dividing the total number of occurrences by the number of participants in the group. For example, if the comparison group used *in my opinion* 20 times, the 20 occurrences were divided by 13 (i.e., there were 13 participants in that group) to arrive at an average number. That number was then used to make comparisons with the other groups' use of the same TFL. Table 8 illustrates descriptive statistics of used formulaic language. *Users* means the number of users who used the formulaic language, and *counts* indicates frequency of the formulaic language occurrences.

Analyzing the frequency and the variety of TFL showed that the learners were not using this formulaic language in their speech initially. Furthermore, both the input enhancement and input + peer-check groups used the TFL more frequently than the comparison group toward the end of the semester. Student A from input + peer-check group did not use TLF to give reasons at pretest (Excerpt 1) but she was able to give clearer reasons by using *one reason is…* and *another reason is …* at posttest (Excerpt 2)

Excerpt 1: Student A from input + peer-check group at pretest

1. *I agree :: with doing club activities is a good idea for students*
2. *I have join a club activity before*
3. *I learned from my experiences*
4. *First I think :: that I can make friends*
5. *And I know about my friends' university and classes*
6. *And we help each other*

Excerpt 2: Student A from input + peer-check group at posttest

1. *In my opinion learning English is important for me*
2. *<u>One reason is</u> :: English is very useful language*
3. *So I can communicate with many foreign people and*
4. *<u>Another reason is</u> :: I can learn other culture*
5. *I know other culture's good points and bad points :: and I can know Japanese culture's good points and bad points*

Moreover, the input + peer-check group was able to produce a wider variety of the TFL than the input enhancement group (Table 8). For example, the input + peer-check group was able to express reasons by saying *It's mainly because….*, *One reason is…*, and *Another reason is….* Conversely, the input enhancement group used one phrase, *It's mainly because….*, to give reasons; this finding indicates that the input + peer-check group had been exposed to or pressured to use a greater variety of phrases through the peer-check intervention, which might have led to a greater improvement in the MLR and PTR.

TABLE 8
Descriptive Statistics of Types of Formulaic Language Used

Group		Opinion					
		In my opinion		Personally speaking, I think		I am not sure, but I think	
Group		pretest	posttest	pretest	posttest	pretest	posttest
Comparison	Users	3	2	0	0	0	0
(*n* = 12)	Counts	3	2	0	0	0	0
	M	0.23	0.15	0	0	0	0
Input enhancement	Users	2	11	0	0	0	1
(*n* = 13)	Counts	2	18	0	0	0	1
	M	0.14	1.29	0	0	0	0.07
+ Peer feedback	Users	4	19	0	4	3	4
(*n* = 21)	Counts	4	20	0	5	3	4
	M	0.19	0.95	0	0.24	0.14	0.19
		Reason					
		It's mainly because		One reason is		Another reason is	
Comparison	Users	0	5	0	0	0	0
(*n* = 12)	Count	0	10	0	0	0	0
	M	0	0.77	0	0	0	0
Input enhancement	Users	0	9	0	0	0	0
(*n* = 13)	Count	0	18	0	0	0	0
	M	0	1.29	0	0	0	0
+ Peer feedback	Users	0	15	0	7	0	7
(*n* = 21)	Count	0	19	0	7	0	7
??	*M*	0	0.9	0	0.33	0	0.33
		Example					
		For example		One example is		Another example is	
Comparison	Users	0	2	0	0	0	0
(*n* = 12)	Counts	0	2	0	0	0	0
	M	0	0.15	0	0	0	0
Input enhancement	Users	2	8	0	1	0	1
(*n* = 13)	Counts	2	10	0	1	0	1
	M	0.14	0.71	0	0.07	0	0.07
+ Peer feedback	Users	4	15	0	0	0	0
(*n* = 21)	Counts	4	23	0	0	0	0
	M	0.19	1.1	0	0	0	0

Discussion

This study examined the development of oral fluency through pedagogical interventions (3/2/1 task, input enhancement, and peer-check for formulaic language). However, not all groups improved in fluency after 13 weeks of implementing the 4/3/2 task. For example, the comparison group which only did the 3/2/1 task showed significantly increased pauses and MDS, thus indicating disfluency. This result differs from those of previous studies, which found that multiple sessions of 4/3/2 training effectively improved fluency (DeJong & Perfetti, 2011).

In contrast, the two intervention groups improved fluency more than the comparison group, with results indicating that using input enhancement and the peer-check of formulaic language led to greater MLR during the academic semester. The input enhancement group gained 15.94% (Cohen's d = .79), and the input + peer-check group gained 30.13% (Cohen's d = 1.31) in MLR, while the comparison group did not show any significant increase (Table 6); this suggests that formulaic language instruction during the 3/2/1 task helped learners improve their MLR, possibly because speakers could access prefabricated chunks stored in long-term memory (Boers & Lindstromberg, 2012; Segalowitz, 2003; Tavakoli et al., 2016; Wood, 2009). With experience and extensive practice, the speed of memory retrieval exceeds that of rule-based processing, and formulaic language is accessed in memory as one unit (Kormos, 2006). In this sense, the participants might have memorized and used certain instances of formulaic language to achieve a particular communicative function.

Input enhancement using the teacher-modeled passage might have helped the participants allocate attentional resources to monitoring how to express their ideas. By reading the model input, the participants might have noticed some of the TFLs because of the underlining cues (Doughty, 1991; Sharwood-Smith, 1993). The TFL can function as discourse markers when organizing an opinion-based monologue. For EFL learners with limited abilities to use English for communication (Leeming & Harris, 2021), arguably knowing how to organize their monologue was a major challenge; therefore, they need function phrases on which to scaffold their ideas. This suggests that using the formulaic language could be an essential device for L2 learners (Zipagan & Lee, 2018).

While input enhancement can raise awareness of noticing the target phrases (Doughty, 1991; Sharwood-Smith, 1993; Szudarski & Carter, 2016), it might not guarantee that the speakers use them during the 3/2/1 task. Conversely, the peer-check might have given the speakers facilitative pressure to use several TFLs. The peer-check function also acted as a peer-feedback of TFL usage, and speakers knew which TFL they did or did not use. Speakers in the input + peer-check group struggled to attend to the TFLs and to the content of their talk during early 3/2/1 training stages; this issue likely arose because attentional resources are limited (Skehan, 1998). Therefore, attention cannot be simultaneously placed on linguistic form and meaning. Gradually, the participants became more at ease using the TFL during the 3/2/1 task.

Second, the previous descriptive analysis (Table 8) suggests that the frequency of the TFL usage is related to the increase of the MLR. This might be because the TFL is usually stored as a whole, which enables the speakers to retrieve these forms faster. Moreover, prefabricated chunks help them to produce longer runs compared to word-by-word retrieval.

Nevertheless, the two experimental groups also produced longer pauses. The participants in both experimental groups in this study increased the MLR while increasing the MLP, indicating that they paused longer when producing longer runs. Given that Towell et al. (1996) have suggested that participants such as those in this study might not have proceduralized their explicit knowledge because their pause length did not decrease, this issue was investigated in depth by analyzing the individual participants' pausing behavior in the 3/2/1 task.

To further discuss which group improved MLR the most, the MLR gain scores (MLR at posttest − MLR at pretest = MLR gain) were converted into *z*-scores. The top six participants' *z*-scores were more than one standard deviation above the mean: Participant 1 ($z = 2.67$), Participant 2 ($z = 1.86$), Participant 3 ($z = 1.80$), Participant 4 ($z = 1.61$), Participant 5 ($z = 1.42$), and Participant 6 ($z = 1.36$) (See Table 9). Five of these participants were from the input + peer-check group. Furthermore, these participants used more TFL at posttest than at pretest, while other participants who did not improve MLR did not use the TFL at posttest.

TABLE 9
Mean Length of Run (MRL) Gains and Usage of the Target Form

Participant	*Group*	*MRL gains*	*Target form frequency (pretest)*	*Target form frequency (posttest)*	*Frequency gain*	*z-score*
1	+ peer feedback	+3.41	3	8	+5	2.68
2	Input enhancement	+2.61	1	1	0	1.86
3	+ peer feedback	+2.54	1	6	+5	1.80
4	+ peer feedback	+2.36	3	9	+6	1.61
5	+ peer feedback	+2.17	1	7	+6	1.42
6	+ peer feedback	+2.11	1	8	+7	1.36

Note. + peer feedback means input + peer feedback group.

Table 9 shows that the participants' MLR gains are somewhat related to their use of the TFL. There are three possible reasons why the input + peer-check group increased both MLR and MLP simultaneously. First, participants paused to concentrate before using the TFL. For instance, Excerpt 4 shows that Participant 6 paused for 1.33 seconds before saying the target form *for example* at posttest:

Excerpt 4: Participant 6, posttest

"It's mainly because (0.43) I can learn many things ***(1.33) <eh> for example*** *(1.49) {I} (1.31) I like Western music and culture of foreign countries"*

Participant 6 had already given a reason why she thought studying abroad was a good idea for university students. She continued to support her opinion by giving an example after pausing. It is plausible that Participant 6 paused to think of the formulaic language and/or an example.

The second reason for the simultaneous increase is that the participant spent time planning after producing the TFL, which sometimes functioned as a filler to give speakers time to think about what to say next. For instance, Excerpt 5 shows that Participant 5 paused for 3.76 seconds after saying *for example*:

Excerpt 5: Participant 5, posttest

<eh> ***for example (3.76)*** *<eh> {you can} (0.36) {you can} (2.26) <eh> you can see the (0.40) movie by you tube (0.36).*

After stating *for example*, Participant 5 reformulated his utterance and repeated *you can* three times, which suggests that he accessed both the conceptualizer and formulator in Levelt's speech model (Levelt, 1989) as he was trying to think of both what and how to say it. Excerpt 6 indicates that Participant 1 paused for 1.91 seconds after saying *in my opinion*:

Excerpt 6: Participant 1, posttest

{in my opinion} *(1.91) <eh> in my opinion learning English is important for me (0.77)*

The speaker had to repair using self-repetition when he said *in my opinion* twice. Repeating the phrase suggested that he was thinking of what to say. That might be why he also spent time accessing appropriate lexis to express himself. Incidentally, online planning time, as indicated by longer pauses, is plausibly needed when learners speak (Yuan & Ellis, 2003).

Considering that this course's objective was to study formulaic language and improve fluency (regardless of assigned conditions), even participants in the comparison group might feel that they need to learn TFL. Therefore, the comparison group differed from control groups in previous studies (Boers, 2014), likely because other researchers did not encourage participants to learn new formulaic language and develop fluency.

However, comparison group participants were rarely observed to transfer what they had learned outside the treatment stage to monologue tasks (Table 8). Linking results and transferring appropriate processing is possible: memory retrieval that L2 learners use during acquisition can best be transferred to a testing situation with characteristics similar to the original learning task (Morris et al., 1977).

Pedagogical Implications

As one of the first attempts to examine formulaic language instruction operationalized during the 4/3/2 task in the regular EFL classroom, this study's findings provide pedagogical implications that are applicable to EFL classrooms. Drawing L2 learners' attention to TFL in the input and encouraging them to practice might impact longitudinal development of their oral fluency within one academic semester.

Awareness-raising activities should be considered to maximize the input enhancement effects in the production stage. Focusing on a given instance of formulaic language once or twice is not enough to leave durable memory traces (Boers et al., 2006); this means that L2 learners can uptake formulaic language through input flooding or ensuring that the same sequence recurs several times in a relatively short stretch

of discourse (Boers et al., 2006). Peer-check activity can help the participants pay attention to a wider range of TFLs, which also leads to improvements in oral fluency.

The 4/3/2 task was originally created as fluency-promoting work; therefore, additional remedial activity might seem to defeat its purpose because L2 speakers need to attend TLFs while they talk. However, as shown in Tran and Saito (2021), this study also revealed that including remedial activity during the 4/3/2 task enhanced fluency.

Conclusions

This study investigated the effectiveness of a fluency-promoting intervention on the longitudinal development of Japanese university students' L2 English-speaking fluency. The study findings supported previous research results indicating that formulaic language helps L2 learners develop oral fluency (Tavakoli & Uchihara, 2020; Tavakoli & Hunter, 2018; Towell et al., 1996; Wood, 2009, 2010, 2015; Wray, 2002). Although the 4/3/2 task was an effective method to enhance oral fluency in previous studies (Bui, 2021; Boers, 2014; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016, Tran & Saito, 2021), additional pedagogical intervention to facilitate the use of formulaic language can enhance fluency development.

This study had several limitations that could have affected the results. First, the groups differed as regards English proficiency. Three groups in this study were formed whose TOEIC scores did not differ appreciably. However, it was not possible to make all groups equivalent concerning English-speaking proficiency in intact classes. The comparison group had significantly better scores on MLR and PTR than the input + feedback group before they started the treatment. Therefore, this group did not perform as hypothesized show much improvement at the end of the academic semester, as was also reported by Towell et al. (1996). Second, there were only 10 TFLs, which would be difficult to generalize to other types of formulaic language. Lower proficiency speakers are likely to rely on formulaic language in a limited manner (Zipagan & Lee, 2018). More types of formulaic language should be examined for future studies. Third, this study did not employ fluency measures to indicate location of pauses (e.g., mid-clause pauses) and perceived fluency rated by human judges (Suzuki & Kormos, 2020). For in-depth understanding of the relationship between frequency usage of formulaic language and oral fluency, future studies should employ additional fluency measurements. Lastly, it should be noted that the analysis in this study did not consider non-fluency-related qualities of the learners' discourse. It remained unclear that use of target phrases may impact on the communicative adequacy, accuracy, and complexity.

Despite these limitations, the current study's findings provide important pedagogical implications for EFL learners' speaking fluency. Research using intact classrooms can increase our understanding of L2 speakers' fluency development inside of the regular classroom. According to Foster's (2020) call for a research agenda focusing on L2 oral fluency pedagogy for the next 10 years, this study can strengthen the connection between research and classroom practice; this may provide classroom teachers insight into fluency-promoting interventions for EFL learners' fluency development.

Acknowledgements

This paper includes data from my doctoral dissertation. I would like to thank Dr. David Beglar for his comments on earlier drafts of this article.

The Author

Chie Ogawa is associate professor in the Dept. of Intercultural Studies of Kyoto Sangyo University in Kyoto, Japan. Her research interests are speaking assessment, interactional competence, Task-based Language Teaching and teacher development.

Department of Intercultural Studies
Faculty of Cultural Studies
Kyoto Sangyo University
Motoyama, Kamigamo, Kita-ku, Kyoto, 603-8555, Japan
Tel: +81 757051411
Email: chieog@cc.kyoto-su.ac.jp

References

Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal, 45*(3), 221-235. https://doi.org/10.1177/0033688214546964

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32*, 83-110. https://doi.org/10.1017/S0267190512000050

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*(3), 245-261. https://doi.org/10.1191/1362168806lr195oa

Boersma, P., & Weenink, D. (2009). PRAAT: Doing phonetics by computer [Computer Software]. http://www.praat.org

Bosker, H. R., Pinget, A., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing, 30*(2), 159-175. https://doi.org/10.1177/0265532212455394

Bui, T. N. (2020). Revisiting the impact of 4/3/2 and 3/3/3 tasks on learners' speaking performance and development: A learner's perspective. *The Journal of Asia TEFL*, *17*(4), 1515-1523. http://dx.doi.org/10.18823/asiatefl.2020.17.4.28.1515

De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of the 6th Workshop on Disfluency in Spontaneous Speech (DiSS)* (pp. 17-20). Stockholm, Sweden: Royal Institute of Technology.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533-568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

De Jong, N. H., Hulstijn, J. H., Schoonen, R., & Groenhout, R. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics, 36*(2), 223-243. https://doi.org/10.1017/S0142716413000210

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*(4), 431-469. https://doi.org/10.1017/S0272263100010287

Dörnyei, Z., & Thurrell, S. (1994). Teaching conversation skills intensively: Course content and rationale. *ELT Journal, 48*(1), 40-49. https://doi.org/10.1093/elt/48.1.40

Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching, 53*(4), 446-461. https://doi.org/10.1017/S026144482000018X

Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*(3), 473-492. https://doi.org/10.2307/3587290

Kehe, D., & Kehe, P. D. (1994). *Conversation strategies*. Pro Lingua.

Kormos, J. (2006). *Speech production and second language acquisition*. Erlbaum.

Leeming, P., & Harris, J. (2020). Expanding on Willis' TBL framework: The integrated input output framework. *The Journal of Asia TEFL*, *17*(1), 215-224. http://dx.doi.org/10.18823/asiatefl.2020.17.1.13.215

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

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

Nation, I. S. P. (1989). Improving speaking fluency. *System, 17*(3), 377-384. https://doi.org/10.1016/ 0346-251X(89)90010-9

Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1*(1), 2-13. https://doi.org/10.2167/illt039.0

Morris, C., Bransford, J., & Franks, J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*(5), 519-533. https://doi.org/10.1016/S0022-5371(77)80016-9

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. https://doi.org/10.1111/lang.12079

Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *Canadian Modern Language Review, 66*(4), 583-606.

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Blackwell.

Segalowitz, N. (2010). *The cognitive bases of second language fluency* (1st ed.). Routledge. https://doi.org/10.4324/9780203851357

Sharwood-Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165-179. https://doi.org/10.1017/S0272263100011943

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

Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, *42*(1), 143-167. https://doi.org/10.1017/S0272263119000421

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics, 26*(2), 245-265. https://doi.org/10.1111/ijal.12092

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task-performance in a second language* (pp. 239-273). Benjamins.

Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly, 50*(2), 447-471. https://doi.org/10.1002/tesq.244

Tavakoli, P., & Hunter, A. M. (2018). Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research, 22*(3), 330-349. https://doi.org/10.1177/1362168817708462

Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning, 70*(2), 506-547. https://doi.org/10.1111/lang.12384

Tran, M., & Saito, K. (2021). Effects of the 4/3/2 activity revisited: Extending Boers (2014) and Thai & Boers (2016). *Language Teaching Research.* https://doi.org/10.1177/1362168821994136

Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly, 50*(2), 369-393. https://doi.org/10.1002/tesq.232

Thomson, H. (2017). Building speaking fluency with multiword expression. *TESL Canada Journal, 34*(3), 26-53. https://doi.org/10.18806/tesl.v34i3.1272

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84-119. https://doi.org/10.1093/applin/17.1.84

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics, 12*(1), 39-57.

Wood, D. (2010). *Formulaic language and second language speech fluency.* Continuum.

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. Bloomsbury.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

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), 1-27. https://doi.org/10.1093/applin/24.1.1

Zipagan, M. N., & Lee, K. R. (2018). Korean English learners' use of lexical bundles in speaking. *The Journal of Asia TEFL*, *15*(2), 276-291. http://dx.doi.org/10.18823/asiatefl.2018.15.2.2.276

(Received August 28, 2021; Revised November 20, 2021; Accepted December 18, 2021)

Appendix A

List of Formulaic Phrases Taught in the Discussion Class

Opinions	
● Asking for Opinions	● Giving Opinions
● What's your opinion?	● In my opinion,…
● What do you think?	● Personally speaking, I think…
● What does everyone think?	● I'm not sure, but I think…
Reasons	
● Asking for Reasons	● Giving Reasons
● Why do you think so?	● It's mainly because…
● How come?	● One reason is…
● Can you tell me why?	● Another reason is…
Examples	
● Asking for Examples	● Giving Examples
● For example?	● For example / For instance,…
● For instance?	● One example is…
● Can you give me an example?	● Another example is…
Joining a Discussion	
● Joining a Discussion	● Asking Others to Join a Discussion
● Can I start?	● Who would like to start?
● Can I say something?	● Does anyone want to comment?
● Can I ask a question?	● Does anyone want to add something?
Possibilities	
● Asking about possibilities	● Talking about possibilities
● If?	● If…
Connecting Ideas	
● Asking Others to Connect	● Connecting to Others' Opinions
● What do you think of {my / name's} idea?	● As {you / name} said,…
● Does anyone agree with {me/ name} ?	

The Journal of Asia TEFL
Vol. 18, No. 4, Winter 2021, 1108-1127

Appendix B

Sample of Input Enhancement

Which is better for you, living in the countryside or living in the city? Why?

In my opinion, it is better to live in an urban area. **It is mainly because** there are more college selections in cities. **For example**, when I was choosing a college near my rural hometown, I had a hard time because I could not find a university that offered what I wanted to study. However, there are many more universities in cities, so you can choose a more suitable one. **Another reason** is that cities offer many opportunities for fun. **For example,** I enjoy going to the theater, and there are many theaters in the city. It is great to be able to see musicals and small theaters live.

Appendix C

Sample of Pair Check Card

Lesson 7
Check card (Opinion, Reason, Example)

	3 minutes	2 minutes	1 minute
Opinion ● In my opinion ● Personally speaking, I think ● I am not sure but I think			
Reason ● It's (mainly/ partly) because ● One reason is ● Another reason is			
Example ● For example/For instance ● One/ Another example is…			