



Effects of the 4/3/2 Activity Revisited: Extending Boers (2014) and Thai & Boers (2016)

Mai Ngoc Tran¹ and Kazuya Saito

Abstract

Whereas scholars have extensively researched on the pedagogical potential of 4/3/2 activities (i.e., increasing time pressure in task repetition) for enhancing L2 fluency, its impact on second language (L2) accuracy remains unclear (e.g., Boers, 2014; Thai & Boers, 2016). To help L2 learners improve *both* fluency and accuracy, the present study examined the effects of adding accuracy enhancement (AE) as a form of delayed metalinguistic correction to the 4/3/2 activity on the development of L2 fluency (speed, breakdown, repair) and accuracy (the acquisition of irregular and regular forms of English past tense). A total of 36 university-level students participated in three 20-minute dyadic sessions. They were randomly divided into three groups: Control, 4/3/2 and 4/3/2+AE. Whereas the 4/3/2 group repeated a monologue task three times with increasing time pressure (4 → 3 → 2 minutes), those in the 4/3/2+AE received metalinguistic correction from the researcher during the 4/3/2 activity. According to the results, a combined approach, 4/3/2+AE, simultaneously impacted learners' overall fluency and accuracy across different topics. However, certain aspects of their fluency and accuracy development, especially those related to linguistic encoding (reduction in pauses within clauses; regular past tense forms), remained unchanged.

Key words: task repetition; corrective feedback; speech production; fluency, accuracy

¹ Corresponding Author: Mai Ngoc Tran, Thu Dau Mot University, maitn@tdmu.edu.vn

Over the past several decades, a growing number of Task-Based Language Teaching studies have examined the facilitative effects of task repetition—a technique in which students will repeat a meaningful task multiple times. A number of scholars have extensively examined the role of increasing time pressure in task repetition (Arevart & Nation, 1991; De Jong & Perfetti, 2011; Gass, Mackey, Fernande, & Alvarez-Torres, 1999; Lambert, Kormos, & Minn, 2016; Maurice, 1983; Nation, 1989). One well-researched repetition is the 4/3/2 activity (Boers, 2014; De Jong & Perfetti, 2011; Maurice, 1983; Thai & Boers, 2016). In this activity, L2 learners are asked to give a monologue on a familiar topic to different partners three times with an increasing amount of time pressure (four, three, and two minutes, respectively). They are paired up with a new interlocutor when they repeat the talk so that they do not feel the need to add new information to it (Arevart & Nation, 1991). If the listener remains the same, the speaker may change the content to keep the listener interested, which will affect their fluency (Boers, 2014).

To date, the 4/3/2 activity has been widely used in many L2 English classrooms worldwide, since the activity is believed to help learners express their ideas more quickly, more efficiently, and with fewer hesitations and shorter pauses. In general, it has been shown that the 4/3/2 activity exerts sizable positive effects on L2 learners' fluency enhancement (e.g., Arevart & Nation, 1991; De Jong & Perfetti, 2011; Nation, 1989). More recently, Boers's (Boers, 2014; Thai & Boers, 2016) work has further looked at the effectiveness of the 4/3/2 activity on various dimensions of L2 speech, finding that the gains were limited to fluent but not to accurate use of language. Motivated by this line of L2 research, the current study set out to further examine how to maximize the pedagogical potential of the 4/3/2 activity. In particular, we corroborated on the extent to which adding accuracy enhancement as a form of delayed corrective feedback (CF) can help Vietnamese English-as-a-Foreign-Language learners improve both fluency and accuracy aspects of their L2 speech.

BACKGROUND

Task Repetition and 4/3/2 as Fluency Enhancement

Task repetition, “the repetition of the same or slightly altered task – whether the whole task, or parts of a task” (Bygate & Samuda, 2005: 43), facilitates students’ L2 speech development. A great deal of attention has been directed toward the content or procedure of task repetition. In some studies, students repeated the same task procedure but with different content (Pinter, 2005). In other studies, students repeated the same task in terms of both content and procedure (Ahmadian & Tavakoli, 2010; Boers, 2014; Lambert et al., 2016; Nation, 1989; Thai & Boers, 2016). Some other studies compared the impact of these different types of task repetition (De Jong & Perfetti, 2011; Gass et al., 1999). Overall, the previous studies have shown that task repetition is particularly effective for developing L2 fluency (Bygate, 2001).

The beneficial effects of task repetition have been explained by many researchers with reference to Levelt’s (1989) model of speech production (e.g., Ahmadian & Tavakoli, 2010). According to Levelt’s model, when speakers intend to speak, they undergo a process comprising of three different stages: conceptualization, formulation, and articulation. In the first stage, learners plan what to say. In the second stage, they engage in lexical, grammatical and phonological encoding to convey the intended message. In the last stage, they turn the planned speech into production. Bygate (2001) asserts that task repetition assists language performance precisely because “part of the work of conceptualization, formulation, and articulation carried out on the first occasion is kept in the learners’ memory store and can be reused on the second occasion” (p. 29).

The 4/3/2 activity is one unique kind of task repetition, whereby students repeat a monologue task three times, but with an increasing amount of time pressure (4 → 3 → 2 minutes). According to Nation (1989), this activity has three important pedagogical features:

- Students repeat the same monologue three times, which reduces the need for planning and promotes their confidence in speaking; and
- The time allowed is reduced each time the talk is delivered, which increases the time pressure believed to support fluency *enhancement*.
- Students are required to repeat the task with different interlocutors.

There have been a number of empirical studies on the effectiveness of the 4/3/2 activity on L2 speech learning (e.g., Arevart & Nation, 1991; Boers, 2014; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016). First, Nation (1989) conducted a study to examine the complexity, accuracy, and fluency aspects of L2 speech among six advanced adult ESL learners who engaged in the 4/3/2 activity. Comparing the participants' first and third speeches, Nation found significant gains especially in their fluency. Arevart and Nation (1991) replicated Nation's (1989) study with 20 intermediate ESL learners in New Zealand. The results of their study were in agreement with those of the original study in indicating that the 4/3/2 technique allowed learners to reach a higher level of fluency. More specifically, in the third delivery of their talks, the number of hesitations reduced significantly. In addition, their speech rate was much faster. However, both Nation (1989) and Arevart and Nation (1991) reported the effects of task repetition within the 4/3/2 sequence. It is worth emphasizing that these two studies made no attempt to examine the long-term effects of task repetition and time pressure.

More recently, De Jong & Perfetti (2011) illustrated in depth the relationship between the 4/3/2 activity and L2 fluency development, controlling for the content of task repetition as one independent variable. In their study, 24 ESL students were divided into two groups. Whereas the repetition group gave a monologue on the same topic three times during the 4/3/2 activity, the non-repetition group talked about three different topics during each delivery. According to the results, the repetition group not only significantly improved their fluency over time, but also successfully transferred such gains to the contexts of new topics. In contrast, the non-repetition group's fluency was clearly observed only when their performance was tested via trained topics.

Particularly relevant to the current study is Boers's research (Boers, 2014; Thai & Boers, 2016). In these seminal studies, the researchers focused on the importance of time pressure by comparing the differential impact of the increasing pressure (4/3/2) and constant pressure (3/3/3) activities on L2 speech development. Interestingly, whereas the 4/3/2 group demonstrated more robust gains in their fluency than the 3/3/3 group did, no significant change was observed among the participants' accuracy and complexity aspects of L2 speech. The results suggest that the increasing time pressure (i.e., 4/3/2) could be instrumental to the development of L2 fluency rather than L2 accuracy, since the time pressure variable induces L2 learners to prioritize fluency over accuracy in L2 speech.

Little is known about what kinds of remedial techniques can best help L2 learners improve both fluency and accuracy when doing the 4/3/2 activity. More research is needed to determine to what degree and how enhancing L2 learners' awareness towards accuracy could be an effective strategy, especially when it is integrated into the context of task repetition. In this regard, the following two studies provide some crucial implications. Hawkes (2011) introduced conscious-raising activities before task repetition. Forty-eight Japanese students with one year's experience of studying English at junior high school participated in this study. According to the results of the interviews, it was shown that these students were guided to pay attention to their accurate use of pronunciation, grammar and vocabulary forms while using language for meaning.

Van de Guchte, Rijlaarsdam, Braaksma, and Bimmel's (2016) evaluated the effects of repeating speaking performance after learners' attention were directed to form during the main task. In this experimental study, a total of 48 Dutch participants learning German as a foreign language first received explicit metalinguistic information on the accurate use of target grammatical structures (the German dative case after a preposition and the German comparatives). Subsequently, the participants were divided into two groups (Task Repetition, No Repetition). Their pre- and post-test performance was then assessed via a range of outcome measures of metalinguistic knowledge and written/oral production. According to the results, although there were no significant differences in their oral performance, the repetition group outperformed the non-repetition group especially in terms of written accuracy. This suggests that the awareness activity can be successful in drawing learners' attention to form, especially when their performance is elicited on written modes, wherein they are allowed to reflect on what they have learned from prior awareness activities.

Taken together, the aforementioned literature provides four broad conclusions. First, repeating the same task multiple times significantly impacts L2 learners' fluency in particular (e.g., Bygate, 2001; Lambert et al., 2016). Second, if such task repetition is carried out under increasing time pressure conditions (4/3/2), it could further enhance the transferability, generalizability, and robustness of L2 fluency development (e.g., De Jong & Perfetti, 2011; Thai & Boers, 2016). Third, the 4/3/2 activity benefits fluency but not accuracy aspects of L2 development, since it induces learners to prioritize fluency over accuracy (Boers, 2014; Thai & Boers, 2016). Fourth, there is an indication that some awareness activities may be needed in the 4/3/2 activity in order to lead to both fluency and accuracy development. For these reasons, in this study we introduced corrective feedback (CF) as accuracy enhancement

during the 4/3/2 activity and scrutinized the differential effects of task repetition, time pressure, and CF on L2 fluency and accuracy development.

Corrective Feedback as Accuracy Enhancement

Corrective feedback has been an area of interest to both second language teachers and researchers since the inception of the field of second language acquisition (Lyster, Saito, & Saito, 2013; Nassaji, 2016). The term CF is defined as “the form of responses to learner utterances containing an error” (Ellis, 2006, p. 28). Whereas the main function of CF is to signal that learners have made errors (i.e., negative evidence), different types of CF can include target forms (i.e., positive evidence) with a different degree of explicitness (e.g., less explicit recasts to more explicit metalinguistic information).

From a theoretical perspective, CF is believed to be effective for L2 learning, especially when it is provided during communicatively authentic tasks, since it encourages learners to notice, understand and acquire form while maintaining their primary focus on meaning (Ellis, 2006). The proponents of the Skill Acquisition Theory also argue that communicative focus on form via CF can promote learners’ gradual transition from controlled to automatic use of language (DeKeyser, 2017). To date, there have been ample studies evidencing the effect of CF on L2 accuracy development in the context of grammar learning (e.g., Li, 2010 and Lyster & Saito, 2010 for meta-analytic reviews) and speech learning (e.g., Saito & Lyster, 2012; Lee & Lyster, 2016).

Interestingly, a majority of CF studies have exclusively focused on how providing CF could facilitate L2 grammar and speech learning when learners are involved with communicatively authentic tasks. To our knowledge, however, few studies have tested the efficacy of CF in the context of task repetition. One such exception is Saito and Akiyama’s (2017) longitudinal intervention study. A total of 30 Japanese EFL students engaged in a weekly dyadic conversation task with native speaking interlocutors via a video conferencing tool over the course of 10 weeks. The native speaking interlocutors provided CF, when the Japanese students’ linguistic errors significantly impaired successful understanding. The results showed that the participants enhanced both accuracy and fluency. However, the operationalization of task repetition was loose. Not only was the structure of each task (e.g., topics, turn taking, time pressure, intervals of repetition) uncontrolled, but the focus and frequency of CF was also inconsistent during task competition.

As shown in Task-Based Language Teaching literature, L2 learners can greatly improve their fluency (but not accuracy) through repeating the same task multiple times, especially under increasing time pressure (4, 3, 2 minutes) (e.g., Thai & Boers, 2016). Thus, it would be interesting to delve into the extent to which adding CF (accuracy enhancement) to the 4/3/2 activity (fluency enhancement) could impact L2 fluency *and* accuracy development. Here, we tested this topic by conducting a quasi-experimental study with a pre-and-posttest design.

CURRENT STUDY

In Boers (2014) and Thai and Boers (2016), it was found that the 4/3/2 activity improved only fluency aspects of L2 speech, indicating that some awareness activities may be needed in order to lead to both fluency and accuracy development (Hawkes, 2011; Van de Guchte et al., 2016). Extending Boers's conceptual framework and research design, and incorporating the perspectives of CF research, the primary objective of the current study is to re-examine the relative effectiveness of the 4/3/2 activity with and without accuracy enhancement, operationalized as a form of delayed CF (for methodological details, see subsequent sections), on L2 accuracy and fluency development. The following research question was formulated:

- To what degree are the effects of task repetitions and increasing time pressure within 4/3/2 activity on utterance fluency and accuracy differentiated when combined with delayed corrective feedback?

Participants

This study was conducted at a university in Vietnam. The participants were composed of 36 Vietnamese EFL learners (7 males and 29 females) who volunteered to participate in the study. Their ages ranged from 18 to 21 (*Age* = 20.2 years). All of the participants reported that they had been learning English for at least 7 years prior to the project. None of them had any experience abroad. Not surprisingly, the participants had few opportunities to use English for communicative purposes outside their classrooms. The participants majored in a wide range of fields: education, information technology, history, law, literature and social work and had to take two General English courses per week at the university. Based on

their self-reported TOEIC scores, the participants' English proficiency levels ranged from pre-intermediate (370 out of 990) to intermediate (690) ($M = 493.1$).

Procedure

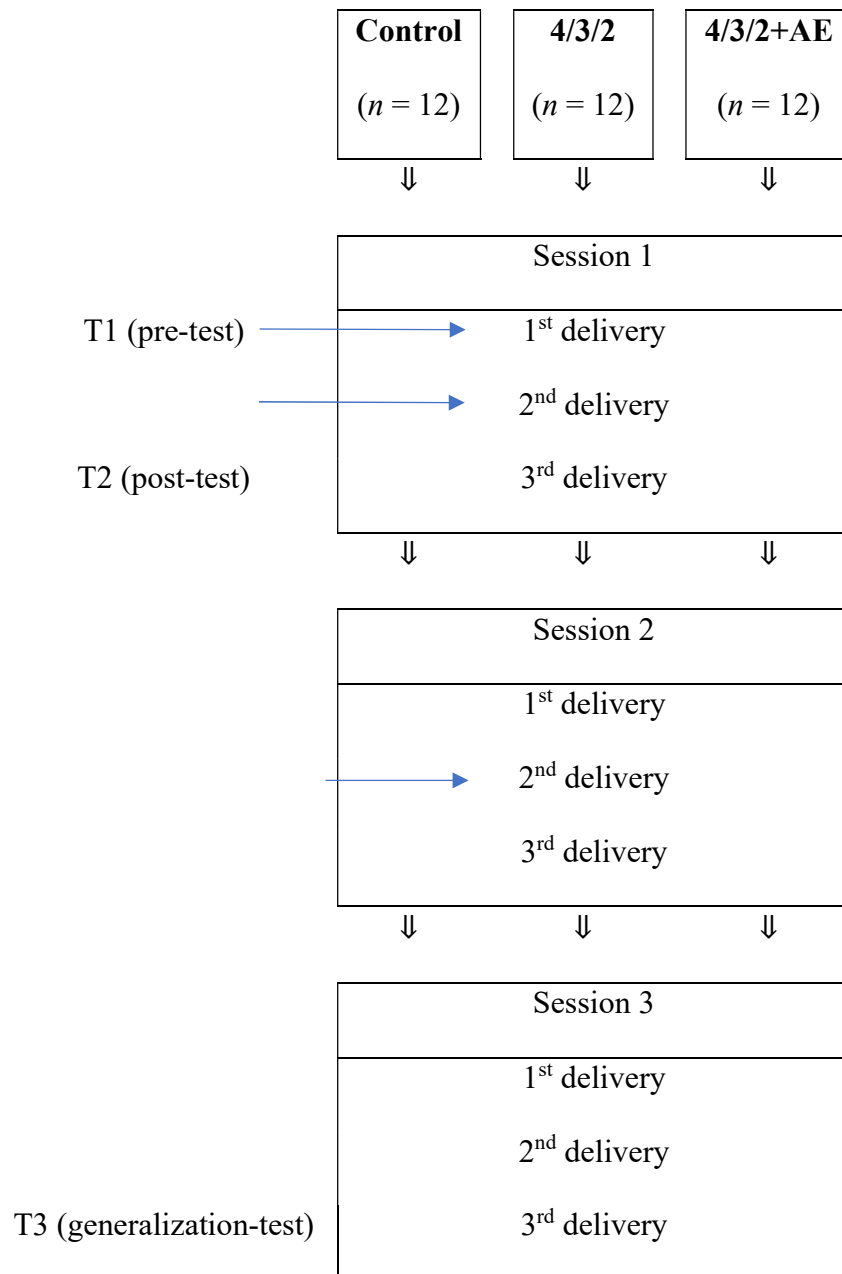
The participants were randomly divided into the three different groups, each of which consisted of 12 students (a) Control, (b) 4/3/2, and (c) 4/3/2 + Accuracy Enhancement (AE). All of them engaged in three individual tutoring sessions (Sessions 1, 2, 3) which took place in a quiet room at the university. Each session lasted for 20 minutes, and the participants repeated a monologue task on one of three different topics (Topics A, B, C: for details see subsequent sections). The interval between these sessions was three days. The nature of instructional treatment was different between the three group conditions as follows:

1. **Control ($n = 12$):** In each session, learners repeated a monologue task three times; a total of three different topics were prepared with prompts to help students organize their talks. The students do not know the topics in advance.
2. **4/3/2 ($n = 12$):** Learners repeated each monologue task three times with increasing time pressure (4, 3, 2 minutes).
3. **Accuracy Enhancement ($n = 12$):** Learners received CF on grammatical accuracy from the researcher between their first and second deliveries during each 4/3/2 activity.

Departing from the previous 4/3/2 literature, which looked at participants' change in fluency and accuracy performance within one single session (e.g., Thai & Boers, 2014), we further aimed to elucidate the extent to which such gains could be generalizable across different sessions. To this end, we tracked participants' fluency and accuracy at three different time points—the 1st and 3rd delivery of the first session (T1, T2) and the final delivery of the third session (T3) (see Figure 1).

FIGURE 1

Design of Current Study



Content of Treatment

In this subsection, we detail the way task repetition, 4/3/2 (increasing time pressure) and accuracy enhancement (corrective feedback) were operationalized.

Task Repetition. All of the participants joined three task repetition sessions. For each session, the following three topics were used in this study: (a) *the last favorite movie you watched*, (b) *your last summer vacation*, and (c) *your 16th birthday celebration* (for the actual materials used in the study, see Appendix). To eliminate any order effects, these topics were used in varied orders. In every group, the first half of the students started with “favorite movie,” then “last summer vacation” and “16th birthday” while the other half started with “last summer vacation,” then “16th birthday” and finally “favorite movie.”

For every session, students first had three minutes of pre-task planning time to organize their talks. To help them conceptualize what to say, they were given six prompts. For example, when students were asked to talk about their favorite movie, the following questions were provided to them: *What was it called? What kind of movie was it? When and where did you watch it? Who were the main characters? What happened in the movie? Why did you like it?* As part of this pre-task planning time, all participants were encouraged to take notes about what they wanted to say by using key words, instead of complete sentences. They were allowed to review their notes only between the first and second deliveries (1 minute) but not between the second and third deliveries. For the same methodological decision, see De Jong and Perfetti (2011) and Thai and Boers (2016).

Increasing Time Pressure (4/3/2). Whereas the participants in the Control group repeated the task in the constant time condition (3 min → 3 min → 3 min), those in the experimental groups (4/3/2, 4/3/2+AE) were given an increasing amount of time pressure for the first delivery (4 min), second delivery (3 min) and third delivery (2 min).

Accuracy Enhancement (Corrective Feedback). To avoid hindering the communicative flow of the participants’ speech, a decision was made to provide delayed CF only between the first and second deliveries, not during their monologue delivery. Recently, a growing number of CF scholars have advocated the effectiveness of delayed CF (e.g., Li, Zhu, & Ellis, 2016; Nassaji, 2011). Different from immediate CF (correcting errors

immediately after they occur), the delayed CF technique allows teachers/researchers to delay the correction until students finish performing their conversations/tasks (Rolin-Ianziti, 2010). To date, certain studies have explored the provision and effectiveness of delayed CF (Nassaji, 2011) or directly compared the different effects of immediate and delayed CF (Li et al., 2016).

Following the delayed CF procedure adopted in Rolin-Ianziti (2010), the researcher wrote down the participants' grammatical errors (past tense) during their first delivery and provided the error logs with metalinguistic explanation to the participants when they finished their monologue. CF was provided in English. In order to make sure these participants clearly understood metalinguistic CF, we kept it as simple as possible. Unlike the participants in the Control and 4/3/2 groups, those who engaged in the accuracy enhancement (delayed CF) did not review their notes on the content. Rather, they reviewed the language-focused feedback from the researcher so that they could reflect on their grammatical accuracy before their second delivery (1 minute).

In terms of the focus of CF, we chose one of the most extensively researched instances: the acquisition of English past tense form (e.g., Li et al., 2016; Yang & Lyster, 2010). This decision was made for the following reasons. First, CF has been found to be effective for the proceduralization of partially acquired knowledge rather than the acquisition of completely new knowledge (Ellis & Sheen, 2006). English past tense is considered to be an ideal testing ground, since all the participants were already familiar with the target structure. English past tense is generally introduced at the very beginning of many textbooks used in Vietnam. Second, despite learners' familiarity and metalinguistic awareness of English past tense, this particular feature has been considered difficult due to the fact that it is perceptibly and communicatively non-salient to many EFL learners (Yang & Lyster, 2010). In fact, many Vietnamese EFL students likely fail to master the past tense during speaking despite many years of exposure to it (McDonald, 2000). Third, the previous CF literature has shown that CF treatment can be effective for the acquisition of the English past tense in many classroom settings (e.g., Yang & Lyster, 2010).²

² It should be noted that in other 4/3/2 studies, students are required to repeat the task with different conversational partners. However, to control the nature and amount of delayed CF treatment, in the present study, the researcher instead of the students' peers acted as their interlocutor. Future studies should further pursue how to train students to become a feedback provider so that both accuracy and fluency activities (tested under the laboratory conditions) can be incorporated to actual classroom settings.

Speech Analyses

Materials. To measure any change or improvement in the students' accuracy and fluency aspects of L2 speech over time, their task performance was timed and audio-recorded using the digital Roland-05 audio recorder (set to 44.1 kHz sampling rate with 16-bit quantization). For the purpose of the analyses, we used the first two minutes from the first and third delivery in the first session (T1, T2); and the first two minutes from the third delivery in the last session (T3) (36 talkers × T1, T2, T3 = 108 samples). The audio-recordings were then transcribed, coded, and analyzed. Fluency measures were analyzed with the aid of the speech analysis program PRAAT (Boersma & Weenik, 2017).

Measuring Fluency. Fluency can be seen as a construct with three subdimensions: breakdown fluency, repair fluency, and speed fluency (Tavakoli & Skehan, 2005; Lambert et al., 2016). Using the fine-grained fluency analysis (looking at speed, breakdown and repair in separation) allows the researcher to see students' speech processes. Following the methodologies that we developed in the precursor fluency studies (e.g., Saito, Ilkan, Magne, Tran, & Suzuki, 2018), the subcomponents of fluency were analyzed as follows:

- **Speed Fluency:** The aspect of speed fluency was assessed in terms of articulation rate by dividing the total number of syllables produced by phonation time. The phonation time was analyzed by subtracting all the fillers (ah, oh, eh) and extensive silence (250ms) from the total length of each sample.
- **Breakdown Fluency:** Pausing behavior was assessed in terms of the ratio of filled and unfilled pauses by dividing the number of pauses by the total number of words. Following Tavakoli and Skehan's (2005) recommendation, we separately calculated breakdown fluency for pauses in the middle and end of clauses. Whereas the ratio of mid-clause pauses is assumed to indicate L2 learners' efficiency in linguistic encoding processes, the ratio of final-clause pauses is supposed to reflect L2 learners' conceptualization processes (Lambert et al., 2016). We followed Bosker, Pinget, Quené, Sanders & de Jong, 2013 and set a *threshold* of 250 ms for pause detection. *Pauses that were equal to or longer than 250 ms were silent (or unfilled) pauses.*

Clause was operationalized “either a simple independent finite clause, or a dependent finite or non-finite clause” (Foster & Skehan 1996, p.310).

- **Repair Fluency:** Repair fluency was analyzed in terms of the ratio of repetitions and self-repairs. This factor is assumed to correspond to L2 learners’ monitoring process (Lambert et al., 2016). The ratio of repetitions and self-repairs was calculated by dividing the number of repetitions and self-repairs by the total number of words. According to Kormos (2000), it is methodologically difficult distinguish between repetitions and self-corrections. While some repeat so that they can secure extra time for monitoring and repair of their production, others repeat as a part of reformulation processes. Following the methodological decisions in the existing literature (Leonard & Shea, 2017; Suzuki & Kormos, 2020), a decision was made to combine both repetition and self-correction measures within a broad category of repair.

Measuring Accuracy. In previous task-based instruction literature, to track changes in accuracy more precisely, the error frequency of a certain specific feature has been analyzed via obligatory context analyses (for the same procedure, see Van de Guchte et al., 2016; Yang & Lyster, 2010). Accordingly, the current study investigated the effects of CF on the participants’ accurate use of the target structure (English past tense) in obligatory contexts. We measured accuracy by dividing the number of target-like forms by the number of obligatory contexts. Furthermore, we used another coding scheme—the acquisition of irregular and regular forms. It has been shown that L2 learners acquire irregular and regular forms at a different rate with the former being mastered more easily and promptly than the latter—a similar developmental pattern in first language acquisition (see Pinker & Ullman, 2002). First, any instances of past tense in obligatory contexts were grouped into either regular or irregular forms. Second, they were coded for Targetlike or Non-targetlike. Table 1 provides some information (e.g., means and standard deviations) regarding raw number of regular and irregular past tense errors.

TABLE 1*Group Means and Standard Deviations for Regular and Irregular Past Tense Errors*

Group	Regular Past Tense						Irregular Past Tense					
	T1		T2		T3		T1		T2		T3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Control	7.50	2.74	8.08	3.91	6.58	3.20	5.25	3.57	6.25	4.26	6.83	4.97
4/3/2	4.67	2.30	6.75	2.59	7.08	2.39	8.33	5.39	10.92	8.20	9.25	6.95
4/3/2+AE	6.25	2.34	7.00	3.54	6.00	2.00	10.67	6.38	4.75	4.22	2.58	1.88

Inter-Coder Reliability. To ensure inter-coder reliability, 20 out of 108 samples (18.5%) were randomly selected and coded independently by a second coder. Two-tailed Pearson correlation analyses demonstrated a high degree of reliability for articulation rate ($r = .955, n = 20, p < .001$), mid-clause pauses ($r = .992, n = 20, p < .001$), final-clause pauses ($r = .983, n = 20, p < .001$), repairs ($r = .811, n = 20, p < .001$), self-repetitions ($r = .904, n = 20, p < .001$), targetlike use of irregular past tense ($r = 1.000, n = 20, p < .001$), non-targetlike use of irregular past tense ($r = 1.000, n = 20, p < .001$), targetlike use of regular past tense ($r = 1.000, n = 20, p < .001$), and non-targetlike use of regular past tense ($r = 1.000, n = 20, p < .001$).

RESULTS

In this section, we aim to examine the extent to which the participants changed their fluency (speed, breakdown, repair) and accuracy (regular, irregular past tense) between T1 and T2 (1st and 3rd deliveries of 1st session) and maintained such change across different topics between T1 and T3 (3rd delivery of 3rd session) according to the three group conditions—Control, 4/3/2, and 4/3/2+AE. Due to the relatively small sample size of the dataset ($n = 36$), we decided to focus only on the participants' improvement over time per group via a set of non-parametric tests *a priori*. Our conservative approach was different from running ANOVAs to find which interaction effects between groups and time points reached significance in a retrospective manner. The alpha level was set at $p < .05$, and adjusted to $p < .008$ via Bonferroni correction. The magnitude of effects was calculated and assessed in conjunction with Plonsky and Oswald's (2014) field-specific benchmarks ($d < 0.70$ for small; $0.70 \leq d < 1.00$ for medium; $1.00 \leq d$ for large effects). According to the results of power analyses via G*Power (Faul, Erdfelder, Buchner, & Lang, 2009), power was .88 with an alpha of .008 and estimated effect size of 0.70, which could be considered as satisfactory in line with the field-specific benchmark (Larson-Hall, 2010).

Fluency Change (T1 to T2) and Generalization (T1 to T3)

According to the results of nonparametric Mann-Whitney tests, prior to the project, the three groups did not significantly differ in articulation rate ($p = .319-.928$), mid-clause pause ratio ($p = .775-.877$), final-clause pause ratio ($p = .119-.443$) and repair ratio ($p = .713-.977$). As summarized in Table 2, the results of nonparametric Wilcoxon Signed

Ranks tests showed that the 4/3/2 group significantly changed speed (articulation rate) and breakdown (mid- and final-clause pause ratio) fluency aspects of language within the first session (T1 to T2) ; and generalized such gains to a different topic context (T1 to T3) with relatively large effects ($d = 0.96-2.35$). Interestingly, although the 4/3/2+AE group's articulation rate and final-clause ratio demonstrated comparable change with large effects ($d = 1.09-2.53$), their performance in mid-clause pause ratio did not reach statistical significance ($p = .028$ for T1-T2, $.041$ for T1-T3). In terms of repair ratio, there was no significant change among any of the three groups ($p > .008$). In short, the results indicated (a) that the 4/3/2 activities could lead participants to attain generalizable change in various dimensions of fluency; and (b) that adding AE to 4/3/2 may negatively impact participants' breakdown fluency within clauses in particular.

TABLE 2*Group Means, Standard Deviations, Significant Change/Generalization and Effect Size for Fluency Performance*

	T1		T2		T3		Change (T1 to T2)			Generalization (T1 to T3)		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>z</i>	<i>p</i>	<i>d</i>	<i>z</i>	<i>p</i>	<i>d</i>
<u>Articulation rate</u>												
Control	2.69	0.34	3.00	0.57	2.84	0.39	-2.35	.019	0.66	-1.17	.239	0.41
4/3/2	2.67	0.26	3.02	0.36	3.13	0.33	-3.05	.002*	1.11	-3.05	.002*	1.54
4/3/2+AE	2.59	0.45	2.86	0.35	3.05	0.39	-2.11	.034	0.66	-2.51	.012*	1.09
<u>Mid-clause pause ratio</u>												
Control	0.20	0.08	0.10	0.06	0.18	0.09	-3.05	.002*	1.41	-.869	.398	0.23
4/3/2	0.21	0.14	0.10	0.07	0.10	0.08	-2.98	.003*	0.99	-2.90	.004*	0.96
4/3/2+AE	0.20	0.10	0.13	0.05	0.12	0.07	-2.91	.028	0.88	-2.04	.041	0.92
<u>Final-clause pause ratio</u>												
Control	0.27	0.06	0.24	0.04	0.21	0.06	-2.11	.034	0.58	-2.27	.023	1.00
4/3/2	0.30	0.06	0.22	0.05	0.18	0.04	-3.05	.002*	1.44	-3.05	.002*	2.35
4/3/2+AE	0.33	0.06	0.23	0.05	0.19	0.05	-2.90	.004*	1.81	-3.05	.002*	2.53
<u>Repair ratio</u>												
Control	.018	.010	.011	.011	.015	.014	-1.95	.050	0.66	-1.02	.308	0.24
4/3/2	.019	.019	.007	.006	.007	.004	-1.68	.091	0.85	-1.88	.059	0.87
4/3/2+AE	.020	.021	.006	.004	.009	.007	-1.96	.050	0.92	-1.72	.084	0.70

Note. * indicates statistical significance, $p < .008$ (Bonferroni corrected)

Accuracy Change (T1 to T2) and Generalization (T1 to T3)

The results of nonparametric Mann-Whitney tests demonstrated the three groups' accuracy performance was comparable at the outset of the project, $p = .590-.713$ for regular past tense accuracy and $p = .152-.443$ for irregular past tense accuracy. The results of nonparametric Wilcoxon Signed Ranks revealed significant change (T1 to T2) and generalization (T1 to T3) among the 4/3/2+AE group's irregular past tense error ratio with large effects ($d = 1.04, 1.42$). The findings here suggest (a) that 4/3/2 only may not impact on L2 accuracy development; (b) that AE may be facilitative of participants' more targetlike use of irregular past tense; and (c) that the role of fluency and accuracy enhancement in the acquisition of regular past tense remains unclear at least in the context of the current study.

TABLE 3*Group Means, Standard Deviations, Significant Change/Generalization and Effect Size for Accuracy Performance*

	T1		T2		T3		Change (T1 to T2)			Generalization (T1 to T3)		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>z</i>	<i>p</i>	<i>d</i>	<i>z</i>	<i>p</i>	<i>d</i>
<u>Regular past tense error ratio</u>												
Control	18.8	18.2	13.1	18.5	28.1	26.5	-1.68	.093	0.31	-0.56	.575	0.40
4/3/2	23.2	22.7	22.4	22.8	20.9	16.0	-0.05	.953	0.03	-.311	.755	0.11
4/3/2+AE	18.4	24.1	21.0	26.5	33.0	13.2	-0.98	.327	0.10	-1.96	.050	0.75
<u>Irregular past tense error ratio</u>												
Control	60.8	23.1	59.0	26.2	54.6	29.5	-0.86	.386	0.07	-1.02	.308	0.23
4/3/2	52.2	27.5	48.0	28.4	54.2	34.1	-1.25	.209	0.15	-0.45	.657	0.06
4/3/2+AE	42.4	42.4	76.2	17.3	86.3	9.7	-3.05	.002*	1.04	-3.06	.002*	1.42

Note. * indicates statistical significance, $p < .008$ (Bonferroni corrected)

DISCUSSION

To date, previous literature has persuasively shown that adding increasing time pressure to task repetition can greatly facilitate L2 fluency change (i.e., 4/3/2 activity). However, the impact of fluency enhancement on L2 accuracy change remains unclear (e.g., Boers, 2014; Thai & Boers, 2016). In this study, we attempted to remedy this issue by incorporating the delayed CF technique (drawing learners' attention for particular linguistic form) as a form of accuracy enhancement (e.g., Li et al., 2016). The results are summarized in Tables 2 and 3. In what follows, we provide an in-depth discussion on the relative effectiveness of the 4/3/2 activity with and without accuracy enhancement.

Increasing Time Pressure

Replicating the previous 4/3/2 literature (e.g., Nation, 1989), we aimed to investigate how to help L2 learners further enhance fluency during task repetition by testing the effect of increasing time pressure—i.e., 4/3/2. According to the results, repeating a task with an increasing amount of time pressure considerably stimulated more fluent speech across both speed and breakdown fluency dimensions to a great degree. However, the participants who received 4/3/2 only did not improve their *accurate* use of L2 English. These findings are in line with previous 4/3/2 studies (Boers, 2014; Nation, 1989; Thai & Boers, 2016). These results were also compatible with those of the online-planning studies (Ahmadian & Tavakoli, 2010). All in all, we echoed these researchers who have argued that when learners are under time pressure to perform a task more rapidly, they are to prioritize meaning conveyance over linguistic accuracy. This relative focus on meaning rather than form may result in L2 fluency, but not accuracy development—i.e., a weakness of the 4/3/2 activity (Boers, 2014; Thai & Boers, 2016).

Adding Accuracy Enhancement to 4/3/2

The current study is one of the first attempts to test the role of accuracy enhancement, operationalized via delayed corrective feedback, in the context of task repetition. In the previous task repetition literature, it has been shown that L2 learners are likely primed to speak fluently rather than accurately during task repetition (Van de Guchte et al., 2016).

Here, we attempted to draw their attention to accuracy via the delayed CF technique, wherein we gave metalinguistic information to participants while they were repeating a task. In this way, we finally tested to what degree and how providing CF during the 4/3/2 activity could impact L2 fluency and accuracy development *simultaneously*. As predicted, the results showed that the 4/3/2+AE group significantly enhanced not only speed and breakdown fluency, but also refine the accurate use of irregular (but not regular) past tense.

In a broader sense, the findings here lend empirical support to the previous task-based instruction studies which emphasized the pedagogical potential of focus on form before students proceed to any types of task repetition activities (Hawkes, 2011; Van de Guchte et al., 2016). In the current study, the participants in the 4/3/2+AE group engaged in the 4/3/2 activities and received delayed CF on English past tense, resulting in fluency and accuracy improvement. On the one hand, the participants were encouraged to enhance all aspects of fluency (speeding up while reducing the number of pauses, repairs and repetitions) thanks to increasing time pressure, as their burden on conceptualization declined through task repetition (Thai & Boers, 2016). On the other hand, delayed CF aided the participants to pay attention to form, even though they concurrently worked on fluency (Li et al., 2016).

In a narrower sense, however, our discussion here (more is better) needs to be considered tentative and interpreted with caution. Notably, the findings presented in this article also hinted not only the potentials of combined accuracy and fluency activities, but also the limits of such composite approach. If we closely look at the details of our fluency and accuracy analyses, the results indeed bring to light the trade-off relationship between fluency and accuracy enhancement and development—i.e., fluency enhancements risks accuracy development (Skehan, 2014).

When comparing the performance of those who received 4/3/2 only versus both 4/3/2 and AE activities, it is important to remember the following slightly different improvement patterns. The 4/3/2-only group greatly developed two dimensions of fluency (speed, breakdown), respectively. In contrast, the 4/3/2+AE group lacked significant improvement in terms of the relatively difficult aspects of L2 fluency (mid-clause pause ratio) and accuracy (regular forms) development (see Tables 2 and 3).

In previous L2 fluency research, this particular dimension of breakdown fluency (mid-clause pause ratio) is thought to relate to L2 learners' linguistic encoding processes (selecting appropriate linguistic forms for the intended message) (Kormos, 2006). Whereas L2 learners quickly reduce the number of pauses at clause-final positions (relevant to

conceptualization), they may need more experience/practice to demonstrate significant, tangible and robust change in their mid-clause pause ratio (e.g., Lambert et al., 2016 for more than three times of repetition; Saito et al 2018 for more than 1 year of immersion experience). When it comes to the L1 acquisition of English irregular vs. regular forms, it has been argued that different processing systems are used with the former being linked to exemplar-based learning and the latter to rule-based learning (Pinker & Ullman, 2002). In the context of L2 learning, there is some evidence (a) that L2 learners of English acquire irregular prior to regular past tense forms in naturalistic (Bardovi-Harlig & Comajoan, 2008); and instructed settings (Yang & Lyster, 2010).

Following this line of thought, we would like to argue that 4/3/2+AE can facilitate L2 learners' overall fluency and accuracy learning only; however, we have yet to know whether 4/3/2+AE could be the most optimal option for promoting all dimensions of L2 speech. When we carefully analyzed the participants' L2 fluency from multiple angles, we did find some evidence of trade-off: Adding accuracy enhancement (CF) seemed to elicit more attention to linguistic accuracy, which could in turn stop learners from reducing mid-clause pauses and enhancing the rule-based learning (the acquisition of regular past tense forms).

Theoretical Interpretations of the Effects of Fluency and Accuracy Enhancement

In terms of utterance fluency, the results showed that the two experimental groups (4/3/2 only, 4/3/2+AE) significantly improved their articulation rate and final-clause pause ratio. Therefore, it seems that the increasing time pressure component might have helped students enhance these two aspects of fluency. Although the 4/3/2+AE group received CF and thus should have assigned attentional resources into form and accuracy aspects to some extent, they could maintain the enhancement of fluency in articulation rate and final-clause pause ratio as much as the 4/3/2-only group did. One explanation for this observation could be that the prompts were so easy that participants may not have needed to process much cognitive resource during conceptualization (particularly on discourse levels). On the other hand, the improvement in mid-clause pause ratio was observed among the control and 4/3/2 only groups. These findings suggest that attention to meaning would enhance the formulation processes which mid-clause pause ratio is supposed to reflect (i.e., disruptions of speech processing) (Kormos, 2006). However, when it comes to the consolidation of formulation processes, it was only the 4/3/2 only group that succeeded in doing so. They sustained the

improvement at the third time point. Those findings indicate that the time pressure in addition to sufficient amount of attention to meaning are necessary for automatization of language-related processing (i.e., formulation).

CONCLUSION, LIMITATIONS AND FUTURE DIRECTIONS

Over the past several decades, L2 researchers and practitioners alike have paid a great amount of attention to task repetition together with some forms of fluency enhancement (i.e., increasing time pressure) (e.g., De Jong & Perfetti, 2011); however, its positive influence on L2 accuracy development remains unclear (e.g., Thai & Boers, 2016). In addition, the role of CF in L2 grammar and speech learning has been one of the most extensively researched areas in the field (e.g., Lyster & Saito, 2010). By interfacing the independently evolving topics (task repetition and CF) in an interdisciplinary manner, the current study scrutinized the complex relationship between the 4/3/2 activity, accuracy enhancement and L2 speech learning.

Overall, the study generated three conclusions. First, the 4/3/2 activity (repeating a monologue task with increasing time pressure) can lead to robust fluency development. Second, drawing learners' attention to form through integrating delayed CF into the 4/3/2 activity can help improve both fluency and accuracy. Third, we have yet to know the extent to which a combined approach (adding AE to 4/3/2) is most effective in helping improve all dimensions of language. When AE is introduced to elicit L2 learners' focus on form, certain aspects of their fluency and accuracy development, especially those related to linguistic encoding (reduction in pauses within clauses), remain unchanged.

To close, we would like to emphasize here again the exploratory nature of the current study. With a view of future replication studies, we acknowledge three crucial limitations in methodology. First of all, this study examined whether CF promoted students' accurate use of the past tense form, a structure that students have already known. In the CF literature, it has been shown that the presence/absence of L2 learners' explicit knowledge for target structures plays a key role in determining the degree of CF effectiveness. According to Ellis & Sheen (2006), CF is more effective if it targets a structure that learners have partial knowledge rather than an entirely new structure. It would be interesting if future studies replicate the findings of the current study but focus on the acquisition of new grammatical features of which learners have no prior knowledge.

Another methodological issue relates to the fact that the current study assessed the students' accuracy based on their targetlike/nontargetlike use of English past tense. The question emerging from this study is whether using global accuracy measures would yield different results (for a list of global accuracy measures, see Yuan & Ellis, 2003; Saito, 2019).

Additionally, although the participants in the current study joined three different sessions of task repetition activities, future studies could adopt a longitudinal design to track L2 learners' performance over multiple repetition sessions for an extensive period of time (e.g., Saito & Akiyama, 2017 for 10 weeks). In this regard, it would be interesting if future studies can highlight both short- and long-term effectiveness of 4/3/2 activities by adopting both immediate and delayed post-tests as done in De Jong and Perfetti (2011). Such studies will shed light on whether the combined effects of task repetition, fluency enhancement and accuracy enhancement will be sustained in the long run and for the sake of the generalizability of the treatment.

In addition, we need to acknowledge that all the task repetition and 4/3/2 activities were operationalized under laboratory conditions with the researcher as a conversational partner, instead of their peers as operationalized in other 4/3/2 studies (Boers, 2014; Thai & Boers, 2016). This decision was made because we wanted to control the nature and amount of delayed CF treatment. However, the generalizability of our results should be tested in classroom settings in the future. Such studies may provide more practical implications for language pedagogy in real classrooms. Also, it should be noted that the current study was conducted with participants whose L2 proficiency levels were considered pre-intermediate and intermediate. The proficiency level of participants may also have affected their development patterns. Therefore, future studies might wish to examine more advanced learners or include different levels of proficiency and treat learner proficiency as an individual learner factor (cf. Lambert et al., 2016).

It is also important to acknowledge that in the present study, only one type of task was used in which six prompts were provided to help students easily plan the structure of speech. Therefore, the cognitive demands on conceptualization could have been relatively low. Future studies could include tasks with higher cognitive demands, which might result in more complex L2 performance (Robinson, 2011, 2015), and for the sake of the generalizability of findings.

Moreover, it should be noted that we combined repetitions and self-repairs based on Lambert, Kormos and Minn's (2016) framework of L2 self-repair. Future research should

analyze ratio of repetitions and self-repairs separately with reference to De Jong, Margarita, Arjen, Rob & Jan (2013) and Kahng (2014) in order to examine the learning and development underlying mechanisms that these two components involve.

Last but not least, one obvious limitation of the current study concerns the relatively small sample size of the dataset, which lowers statistical power in all the statistical analyses. Therefore, future studies with larger sample size are recommended in order to replicate the findings in the present study with a view of obtaining more robust understanding of the complex relationship between accuracy, fluency, and training over time.

REFERENCES

- Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15 (1), 35–59.
- Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal*, 22(1), 84–94.
- Bardovi-Harlig, K., & Comajoan, L. (2008). Order of acquisition and developmental readiness. In B. Spolsky and F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 383–397). Malden, MA: Blackwell.
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal*, 45 (3), 221–235.
- Boersma, D., & Weenink, P. (2017). *Praat: Doing phonetics by computer version 6.0.29*. Retrieved from <http://www.praat.org>
- Bosker, H. R., Pinget, A.-F., 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. Doi: 10.1177/0265532212455394
- 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.
- Bygate, M. & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in second language* (pp. 37–74). Amsterdam: John Benjamins.
- DeKeyser, R. M. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15-32). New York and London: Routledge.
- 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.
- De Jong, N. H., Margarita P. S., Arjen, F. F., Rob, S., & Jan, H. H., 2013. Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics* 34(5):893–916.
- Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review*, 19, 18–41.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28 (4), 575–600.

- 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.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323.
- Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49* (4), 549–581.
- Hawkes, M. L. (2011). Using task repetition to direct learner attention and focus on form. *ELT Journal, 66* (3), 327–336.
- Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning, 64*(4), 809–854.
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, N. J.: Lawrence Erlbaum Associates.
- Kormos, J. (2000). The timing of self-repairs in second language speech production. *Studies in Second Language Acquisition, 22*(2), 145–167.
<https://doi.org/10.1017/S0272263100002011>
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
- Lambert, C., Kormos, J., & Minn, D. (2016). Task repetition and second language speech processing. *Studies in Second Language Acquisition, 39* (1), 167–196.
- Lee, A. H., & Lyster, R. (2016a). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition, 38*, 35-64.
- Leonard, K. R., & Shea, C. E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity, and underlying cognitive factors. *The Modern Language Journal, 101*(1), 179–193. <https://doi.org/10.1111/modl.12382>
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60* (2), 309–365.
- Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal, 100* (1), 276–295.

- Lyster, R., & Saito, K. (2010). Corrective feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265-302.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46, 1-40.
- Maurice, K. (1983). The fluency workshop. *TESOL Newsletter*, 17 (4), 29.
- McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21 (3), 395–423.
- Nassaji, H. (2011). Correcting students' written grammatical errors: The effects of negotiated versus nonnegotiated feedback. *Studies in Second Language Learning and Teaching*, 1 (3), 315–334.
- Nassaji, H. (2016). Anniversary article interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20 (4), 535–562.
- Nation, P. (1989). Improving speaking fluency. *System*, 17 (3), 377–384.
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6 (11), 456-463.
- Pinter, A. (2005). Task repetition with 10-year old children. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113–126). Basingstoke: Palgrave Macmillan.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64 (4), 878–912.
- Robinson, P. (2011). *Second language task complexity: researching the cognition hypothesis of language learning and performance* (Vol. 2). Amsterdam: John Benjamins.
- Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In *Domains and directions in the development of TBLT* (pp. 87–122). John Benjamins Amsterdam. Rolin-Ianziti, J. (2010). The organization of delayed second language correction. *Language Teaching Research*, 14 (2), 183–206.
- Saito, K. (2019). To what extent does long-term foreign language education improve spoken second language lexical proficiency? *TESOL Quarterly*, 53, 82-107.
- Saito, K., Ilkan, M., Magne, V., Tran, M., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low, mid and high-level second language fluency. *Applied Psycholinguistics*, 39, 593-617.

- Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language learning*, 67, 43-74.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning*, 62, 595–633.
- Segalowitz, N. (2012). Second language fluency. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 240–244). New York: Routledge
- Skehan, P. (2014). Limited attentional capacity, second language performance and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–262). Amsterdam: Benjamins.
- 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>
- 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. 239–277). Amsterdam: Benjamins.
- Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, 50 (2), 369–393.
- Van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2016). Focus on form through task repetition in TBLT. *Language Teaching Research*, 20 (3), 300–320.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32 (2), 235–263.
- 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.

APPENDIX

Topics Used in the Three Sessions

Topic A: The Last Favorite Movie You Watched

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. What was it called?
2. What kind of movie was it?
3. When and where did you watch it?
4. Who were the main characters?
5. What happened in the movie?
6. Why did you like it?

Begin your talk with the first sentence given

The favourite movie I recently watched was...

Topic B: Your Last Summer Vacation

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. Did you enjoy your last summer vacation? If so, what did you enjoy about it? If not, why did you encounter any problems?
2. Where did you go?
3. Who did you go with?
4. How did you get there?
5. How long did you stay there?
6. What did you do?

Begin your talk with the first sentence given

For my last summer vacation, I went to....

Topic C: Your 16th Birthday Celebration

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. How did you celebrate your 16th birthday?
2. Where did it take place?
3. Who did you invite?
4. What kinds of presents did you receive?
5. What did you eat?
6. Did you enjoy your party? Why/ Why not?

Begin your talk with the first sentence given

I celebrated my 16th birthday...