Corrective Feedback, Spoken Accuracy and Fluency, and the Trade-off Hypothesis

Mohammad Hassan Chehr Azad
PhD candidate, Corresponding author, University of Tabriz, Iran, *chehrazad88@ms.tabrizu.ac.ir*

Farahman Farrokhi
Prof., University of Tabriz, Iran, *ffarrokhi20@yahoo.co.uk*

Mohammad Zohrabi
Asst. Prof., University of Tabriz, Iran, *mohammadzohrabi@gmail.com*

The current study was an attempt to investigate the effects of different corrective feedback (CF) conditions on Iranian EFL learners' spoken accuracy and fluency (AF) and the trade-off between them. Consequently, four pre-intermediate intact classes were randomly selected as the control, delayed explicit metalinguistic CF, extensive recast, and intensive recast groups. All participants took part in spoken reproduction tasks for six sessions and their spoken grammatical errors were treated differently. Then, the data was transcribed, coded for AF, and analyzed. The results indicated that the different CF conditions had insignificant effects on spoken general accuracy. Considering fluency, there was a significant correlation between pruned and unpruned speech rates, with a large effect size; however, the different CF conditions had insignificant effects on the spoken temporal fluency. The correlations between general accuracy and temporal fluency of all groups were positive and significant with a small effect size. In addition, different CF conditions had insignificant effects on the relationship between accuracy and fluency. These findings suggest that there is no trade-off between general accuracy and temporal fluency. In addition, different CF conditions, particularly when they are focused on a specific form, have insignificant effects on the EFL learners' spoken general accuracy, temporal fluency, and their relationships.

Keywords: corrective feedback (CF), focus on form, spoken general accuracy, spoken temporal fluency, trade-off hypothesis

INTRODUCTION

One of the shortcomings of the meaning focused instruction, which helped the learners become fluent, was its insufficiency to ensure comparable levels of accuracy and,

consequently, this led to the introduction of the focus on form approach (Long, 1991), a subordinate category of form focused instruction. Accordingly, "one of the methodological macro-options for focus on form is corrective feedback options" (Ellis, 2008, p. 869). CF, as an effective way to promote noticing, is considered conducive to L2 learning and grammar development (Ammar & Spada, 2006; Ellis, Loewen, & Erlam, 2006; Li, 2010; Lyster, 2004; Sheen, 2007). It can also develop learners' specific spoken accuracy (Chehr Azad, Farrokhi, & Zohrabi, 2017). However, production isn't a one-dimensional concept which is restricted to accuracy. Indeed, L2 proficiency is a multi-componental construct consisting of three principal dimensions of complexity, accuracy, and fluency (Ellis & Barkhuizen, 2005; Skehan, 1996, 1998; Skehan & Foster, 2001). Consequently, because of attentional limitations, the development of one of these aspects might be at the expense of the others (Skehan, 1998). For example, accuracy development might be at the expense of fluency development because of the learners' attentional limitations. There have been several studies (Ansarin & Chehr Azad, 2015; Farrokhi & Chehr Azad, 2012; Hoseini Fatemi & Harati, 2014; Maftoon & Kolahi, 2009; Salimi, 2015) investigating the effects of the different CF conditions on the EFL learners' spoken accuracy. There have also been some studies (e.g., Chehr Azad, Farrokhi, & Zohrabi, 2017; Rahimi & Vahid Dastjerdi, 2012; Sato & Lyster, 2012; Seyed Motahari & Ghasemi Nik Manesh, 2014) examining the effects of the CF on CAF. However, no studies, to the researchers' knowledge, have studied the effects of the CF on EFL learners' general accuracy and temporal fluency and the trade-off between them. Therefore, the current research was an attempt to accomplish these.

LITERATURE REVIEW

Focus on Form

According to Long (1991), "focus on form overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (pp. 45-46). Doughty (2001) also noted "the factor that distinguishes focus on form from other pedagogical approaches is the requirement that focus on form involves learners briefly and simultaneously attending to form, meaning, and use during one cognitive event" (p. 211).

Corrective Feedback (CF)

According to Ellis (2008), "one of the methodological macro-options for focus on form is CF" (p. 869). Ellis, Loewen, and Erlam, (2006) defined it in the following terms:

> Corrective feedback takes the form of responses to learner utterances that contain an error. These responses can consist of (a) an indication that an error has been committed, (b) provision of the correct target language form, and (c) metalinguistic information about the nature of the error, or any combination of these. (p. 340)

Proponents of noticing hypothesis (Ellis, 1991; Gass & Varonis, 1994; Sato & Lyster, 2012; Schmidt, 1990, 2001) have considered CF as a means of drawing learners' attention to form and as a stimulus for noticing. In addition, it gives them an opportunity

to make a cognitive comparison between their interlanguage and the given input (Ellis, 1994). It can also help them engage in focused input analysis (Ellis, 2005).

Different CF Types

One CF type which is relevant to the current study and which appeals to researchers, due to theoretical and practical reasons (Ellis, 2008), is recast. Lyster and Ranta (1997) defined it as "the teacher's reformulation of all or part of a student's utterance, minus the error" (p. 46). Recasts can be categorized into intensive, focused, and extensive, unfocused, types (Ellis, 2001; Loewen, 2011). According to Ellis (2001), intensive recasts occur when the single target structure is selected in advance, and learners are likely to receive feedback multiple times on it. In other words, "when intensive recasts are provided, errors related to the target structure are the only ones addressed" (Kamiya, 2015, p. 60). In contrast, extensive recasts occur when feedback is not limited to a single target structure and learners receive feedback on many structures that occur incidentally during the instruction. It should also be mentioned that when intensive recasts are provided, the number of recasts focused on a single target structure is likely to be higher than when extensive recasts are provided and they can be considered as an explicit type of feedback (Kamiya, 2015). On the other hand, extensive recasts are directed at different structures, and they can be considered as an implicit type of feedback.

There have been few studies which were based on the simultaneous comparison of the effectiveness of both intensive and extensive recasts. In some of these studies, there were no significant differences between these two types of recasts. For example, a meta-analysis conducted by Russell and Spada (2006) showed that there was no difference between intensive and extensive CF. In addition, Ellis, Sheen, Murakami, and Takashima (2008) compared the effects of intensive and extensive recasts and showed that there was no statistically significant difference between their effects on the development of English articles. Other studies, in contrast, have supported the superiority of intensive recasts. For instance, Nicholas, Lightbown, and Spada (2001), reviewing previous studies, discovered that intensive recasts were more effective than extensive recasts. Furthermore, in their meta-analysis, Mackey and Goo (2007) showed that intensive recasts were more effective than extensive recasts.

Another CF type which is relevant to the current study is called explicit CF. Sheen, Wright, and Moldawa's (2009) study showed the superiority of intensive recasts over extensive recasts. According to Ellis (2008), explicit correction is "an utterance that provides the learner with the correct form while at the same time indicating an error was committed" (p. 227). The following example is taken from Ellis (2009a).

L: On May.

T: Not On May. In May. We say "It will start in May" (p. 9).

The other CF type, relevant to the current study, is metalinguistic CF. According to Ellis (2008), it refers to "an utterance that provides comments, information, or request related to the well-formedness of the learner's utterance" (p. 227).

L: I go to Paris last year.

T: Went. You should use simple past tense.

Aspects of the Spoken Production

According to SLA researchers (Ellis & Barkhuizen, 2005; Skehan, 1996, 1998; Skehan & Foster, 2001), L2 proficiency and L2 performance are multi-componental constructs consisting of three principal dimensions of complexity, accuracy, and fluency (CAF). One aspect of spoken production, which is relevant to the current study, is accuracy. Ellis (2003) referred to accuracy as "the extent to which the language produced in performing a task conforms to the target language norms" (p. 339). Another aspect of spoken production, which is relevant to the current study, is spoken fluency. According to Lennon (2000), fluency is "the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing" (p. 26). Similarly, Skehan (1996) defined it as the learners' capacity to mobilize their interlanguage systems to communicate meanings in real time.

Trade-off Hypothesis

Skehan (1998) came up with his trade-off hypothesis, also known as the limited attentional capacity model, stating that CAF are interdependent such that increased performance in one area may occur at the expense of performance in the other areas. Indeed, this hypothesis predicted that "committing attention to one area, other things being equal, might cause lower performance in others" (Skehan, 1998, p. 112). In particular, it is proposed that there might be a tension between form, complexity and accuracy, and fluency (Skehan, 1998). This tension and prioritization might have some consequences. For example, consistent prioritization of fluency might lead to over-lexicalized performance and fossilized language which may be difficult to change. Consistent prioritization of accuracy, in contrast, might lead to lack of fluency. Therefore, the nature of the trade-offs and balancing these aspects of performance should be thoroughly considered.

Considering the nature of the trade-off, Foster and Skehan (1996) have argued that the trade-off is between accuracy and complexity. Skehan and Foster (1997) also reported a trade-off between accuracy and complexity in a study focusing on the effect of planning during three oral tasks. Finally, Skehan (2009) suggested that accuracy and complexity rarely go together. Other studies, on the other hand, have suggested that the trade-off is between meaning, fluency, and form, either complexity or accuracy. For example, some studies (Ahmadian & Tavakoli, 2011; Wendel, 1997; Yuan & Ellis, 2003) have proposed that the trade-off involves fluency and accuracy. Brumfit (1984), focusing on the fluency-form distinction, also claimed that spontaneous and free-flowing language is the goal of fluency-oriented tasks and a focus on form and control is the goal of accuracy-oriented tasks. Skehan (1998) also stated that "adult learners vary in learning style by learning through exemplars and emphasizing fluency or by learning through analysis and emphasizing complexity or accuracy" (p. 269).

The Present Study

There were two purposes for the current study. The first one was to examine the different effects of different corrective feedback (CF) types on Iranian EFL learners' general accuracy, operationalized as error free clauses, and temporal fluency,

operationalized as pruned and unpruned speech rates. The second one was to examine the presence or absence of a trade-off between accuracy and fluency and how it was affected by different CF types. To attain these, the following research questions were formulated.

Research Question 1(RQ1): Are there any significant differences among different CF types' effects on Iranian EFL learners' spoken general accuracy, error free clauses?

Research question 2 (RQ2): Are there any significant differences among different CF types' effects on Iranian EFL learners' spoken temporal fluency, pruned speech rate?

Research Question 3 (RQ3): Is there a trade-off between Iranian EFL learners' spoken accuracy and fluency (AF) and how is it affected by different CF types?

METHOD

Design of the Study

The current study was quasi-experimental. Its independent variable was focus on form, operationalized as corrective feedback (CF), with four levels of no CF (n = 16), intensive recast (n = 15), extensive recast (n = 16), and delayed explicit metalinguistic feedback (n = 17). The dependent variables of the study were general accuracy, operationalized as error free clauses, and temporal fluency, operationalized as pruned and unpruned speech rate. Four pre-intermediate intact classes were randomly selected and assigned into the control, delayed explicit, extensive recast, and intensive recast groups.

Participants

This study was conducted at a private English language learning school in Tabriz, Iran. Four pre-intermediate intact classes with 64 male learners who were bilingual speakers of Azeri and Persian were randomly selected. They were between the ages of 15 and 23. The course they were taking was based on task-based language teaching. Their weekly attendance at school was 3 sessions. They had no opportunity for informal interaction in English outside the classroom. They were considered a fairly homogenous pre-intermediate group of learners, based on their learning history and English proficiency. To verify their initial homogeneity, a Key English Test (KET) was used and the test results were analyzed via a one-way ANOVA, with the alpha set at .05, and it was revealed that they were initially homogenous ($F_{3,60} = .28$, $p = .87$).

Procedure

One week before the commencement of the study, KET (Key English Test) was administered to ensure their initial homogeneity. In addition, one week later, the first written fill in the blank test, the written pretest, was given to all groups' participants to ensure their initial homogeneity with respect to their knowledge of the simple past tense and it was revealed that they were homogenous ($F_{3,60} = .28$, $p = .88$).

The basic process was carried out during six instructional sessions, each of which was divided into two halves. The first half was based on following the institute's term

program. The other half was devoted to the main process of the current study and was audio recorded for the subsequent analyses. Each of these began by assigning a story from Steps to Understanding (Hill, 1988) to all participants of all groups. Indeed, while within group stories were different due to the prevention of the practice effect, between group stories were the same in each of these instructional sessions. In other words, in each of these sessions, participants within each group were to read a different story, summarize it and retell it to the class. Between group participants, in contrast, were to read the same stories. Based on the previously conducted study (Chehr Azad et al., 2017), all groups' participants were given 5 min and asked to read, summarize and plan their assigned stories. When they completed the task of summarizing and planning of their individualized stories, they were asked to orally reproduce and retell them to the whole class. During the oral reproductions of the stories all groups experienced different processes. Indeed, they were different with respect to the presence or absence of the CF, its focus, its type, and the time it was provided.

The control group and the quasi-experimental groups were different based on the presence or absence of the CF. That is, while the control group's participants received no CF on their errors, the quasi-experimental group's participants received CF on their errors. In addition to the differences between the quasi-experimental groups and the control group which were based on the presence or absence of CF, all quasi-experimental groups were also different based on the focus and types of the CF and the time it was provided. In the intensive (focused) recast group, the CF was immediately and intensively provided on the participants' simple past tense errors during their oral reproductions. In other words, as soon as they made simple past tense errors on their oral reproductions of their stories, they were provided with CF of the recast type. In the extensive (unfocused) recast group, unlike the focused recast group, the recast was extensively provided on all errors of their oral reproductions of the stories. In the delayed explicit metalinguistic corrective feedback group, in contrast, there was no immediate CF and the type of the CF was explicit and metalinguistic. That is, during their story reproductions, the simple past tense errors were not immediately corrected and the CF was provided at the end of their story reproductions. In other words, the researcher did not immediately react to their errors and made notes of their simple past tenses errors during their story reproductions. At the end of each participant's oral reproduction, he presented the simple past tense errors on the board, corrected explicitly, and provided some metalinguistic explanations for all participants of the group. This process of story summarizing, its oral reproduction, and error treatment lasted for six sequential sessions. Then, the recorded oral data of all participants during 6 instructional sessions were transcribed and coded by the researcher focusing on spoken accuracy and fluency.

General Accuracy and Temporal Fluency Measurements

Following previous studies (Foster & Skehan, 1996; Skehan & Foster 1997, 1999, 2005; Vercellotti, 2012; Wendel, 1997; Yuan & Ellis, 2003), the accuracy measurement of the learners' spoken production was based on the proportion of the error free clauses.

The approach which was taken to measure spoken fluency was also based on previous studies (Saeedi, 2015; Sato & Lyster, 2012). In this approach, two different but relevant fluency measurement indices, unpruned and pruned speech rates (SRs), were calculated on the basis of Lennon's (as cited in Sato & Lyster, 2012) suggestion that pruning speech samples gives more reliable scores of fluency. Unpruned speech rates were obtained by dividing the total number of words each learner produced in each of the oral reproduction tasks and sessions of the study by the total delivery time in miliseconds. Then, this number was converted to obtain words-per-minute scores. Pruned SRs were also calculated in the same manner. However, the pauses, repetitions, and hesitation markers were excluded from the word count.

After the first identification of the measurement units and calculation of the different indices of CAF, there was a need to measure intra-rater reliability, the reliability of the researcher's scoring. To this end, to calculate intra-rater reliability, the researcher recoded and recalculated different AF indices in a different order. The intra-rater reliability using Cohen's Kappa was .86. In addition to the measurement of intra-rater reliability, there was a need to measure inter-rater reliability. To this end, another researcher, one of the researcher's colleagues independently coded 15% of the data. The inter-rater reliability using Cohen's Kappa was α = .80.

DATA ANALYSIS

First, the descriptive statistics were calculated. Then, a one-way ANOVA was used to analyze the participants' spoken general accuracy, error free clauses, during six sessions. To analyze the fluency measurements, it was needed to take an additional step. That is, since the temporal fluency consisted of two indices including pruned and unpruned speech rates, it was needed to choose one of these. Consequently, a 2×2 correlation matrix was created with Pearson correlation coefficients, using scores of all participants, to explore the relationships between them and choose one of them. Another one way ANOVA was used to analyze the data for the temporal fluency, pruned speech rate. Next, a 2×2 correlation matrix was created with Pearson correlation coefficients to study the relationship between general accuracy, error free clauses, and fluency, pruned speech rate. Finally, to study the effects of different CF types on the relationships between accuracy and fluency, eight 2×2 correlation matrices were created with Pearson correlation coefficients using each of the participating groups' scores in both sessions 1 and 6.

FINDINGS

As it is depicted, general accuracy of the control group during sessions 1 and 4 unsteadily changed. However, it declined between sessions 4 and 6.
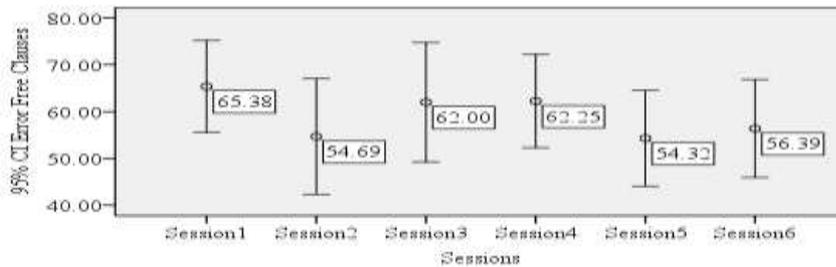
Figure 1
Control group's general accuracy measurement

In addition, the delayed explicit group's spoken general accuracy, error free clauses, between sessions 1 and 6 is given in Figure 2.
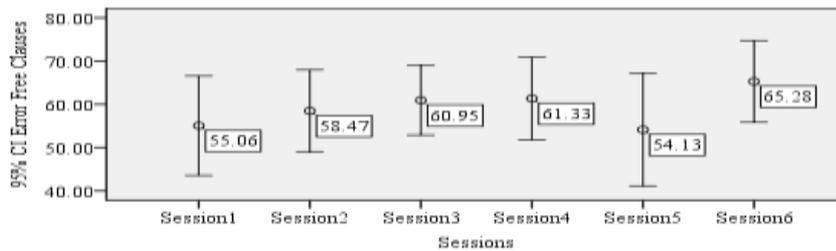

Figure 2
Delayed explicit metalinguistic group's general accuracy measurement

As it is depicted, there was a steady increase in general accuracy of the delayed explicit metalinguistic group between sessions 1 and 6.

Furthermore, the extensive group's spoken general accuracy, error free clauses, between sessions 1 and 6 is given in Figure 3.
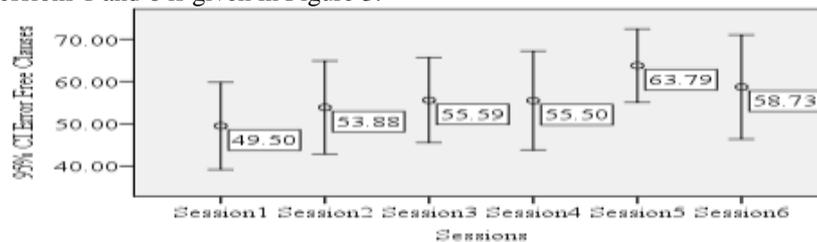

Figure 3
Extensive recast group's general accuracy measurement

As it is shown, there were unsteady rises and falls in general accuracy of the extensive recast groups between sessions 1 and 6.

Finally, the intensive group's spoken general accuracy, error free clauses, between sessions 1 and 6 is given in Figure 4.
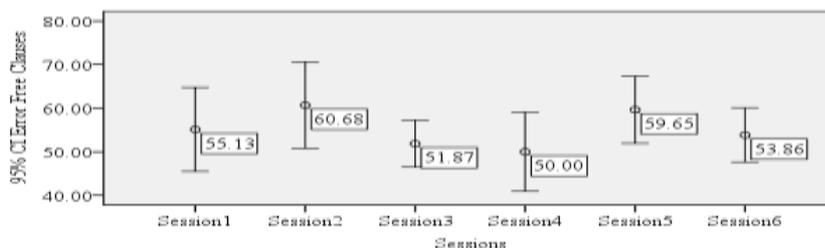
Figure 4
Intensive recast group's general accuracy measurement

As it is shown, there were unsteady rises and falls in general accuracy of the extensive recast groups between sessions 1 and 6.

To summarize, although general accuracy of the control group during sessions 1 and 4 unsteadily changed, it declined between sessions 4 and 6. There were also unsteady rises and falls in general accuracy of both the extensive and intensive recast groups between sessions 1 and 6. However, there was a steady increase in general accuracy of the delayed explicit metalinguistic group between sessions 1 and 6.

Considering the fact that the comparisons were based on sessions 1 and 6, only these sessions' results are comprehensively presented. Consequently, the results of the descriptive statistics of all groups' general spoken accuracy, error free clauses, in sessions 1 and 6 are given in Table 1.

Table 1
Descriptive statistics of groups' general spoken accuracy in sessions 1 and 6

Sessions		*n*	*Mean*	*SD*	*95% Confidence Interval*	
					Lower Bound	Upper Bound
Session 1	Control	16	65.38	18.41	55.57	75.18
	Delayed Explicit	17	55.06	22.40	43.54	66.58
	Extensive Recast	16	49.50	19.41	39.16	59.84
	Intensive Recast	15	55.13	17.38	45.51	64.76
Session 6	Control	18	56.39	21.07	45.91	66.87
	Delayed Explicit	18	65.28	18.92	55.87	74.69
	Extensive Recast	11	58.37	18.37	46.38	71.07
	Intensive Recast	14	53.86	10.77	47.64	60.08

As it is shown, in both the extensive recast and delayed explicit groups, the general accuracy measurement of session 6 was bigger than that of session 1. In both the control and the intensive recast groups, in contrast, it was smaller than that of session 1. In addition, the delayed explicit had the highest general accuracy in session 6.

In addition, the results of the one-way ANOVAs used to analyze the participants' spoken general accuracy in sessions 1 and 6 revealed that there were insignificant differences among all groups in both sessions 1($F_{3, 60} = 1.84$, $p = .16$) and 6 ($F_{3, 60} = 1.24$, $p = .31$). In other words, the effects of different CF conditions on the learners' production of the error free clauses and their general spoken accuracy were insignificant.

Fluency

To choose one of the temporal fluency indices, there was a need to run another correlational analysis. To this end, another 2×2 correlation matrix was created with Pearson correlation coefficients, using the participants' grades on these indices and its results revealed that the correlation between pruned and unpruned speech rate was positive and significant ($r = .87$) and the effect size was large ($R^2 = 0.8$). Since the temporal fluency index based on pruned speech rate was calculated after subtracting repair fluency indices including repetitions, replacements, reformulations, and false starts, it was chosen as the index representing the temporal fluency for further analysis.

First, the control group's spoken temporal fluency, pruned speech rate, between sessions 1 and 6 is given in Figure 5.
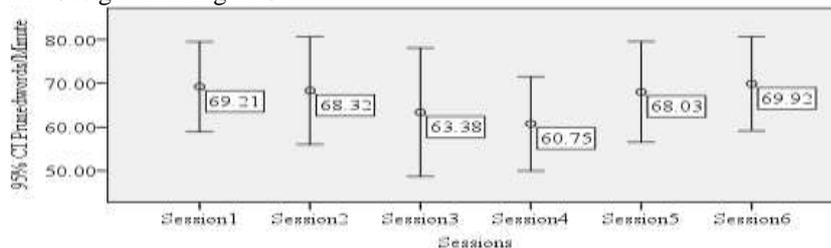

Figure 5
Control group's temporal fluency measurement

As it is depicted, temporal fluency of the control group during sessions 1 and 4 declined. However, it increased between sessions 4 and 6.

In addition, the delayed explicit group's spoken temporal fluency, pruned speech rate, between sessions 1 and 6 is given in Figure 6.
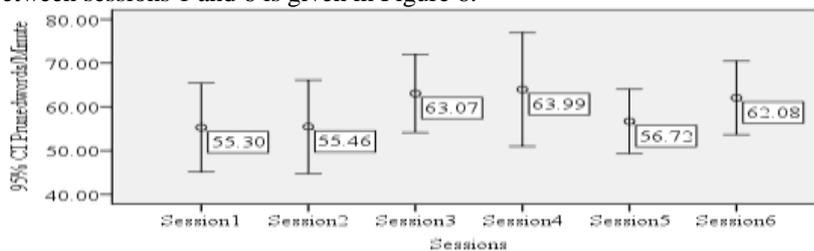

Figure 6
Delayed explicit metalinguistic group's temporal fluency measurement

As it is depicted, there was a steady rise in temporal fluency of the delayed explicit metalinguistic group between sessions 1 and 4, but it unsteadily changed between sessions 4 and 6.

Furthermore, the extensive recast group's spoken temporal fluency, pruned speech rate, between sessions 1 and 6 is given in Figure 7.
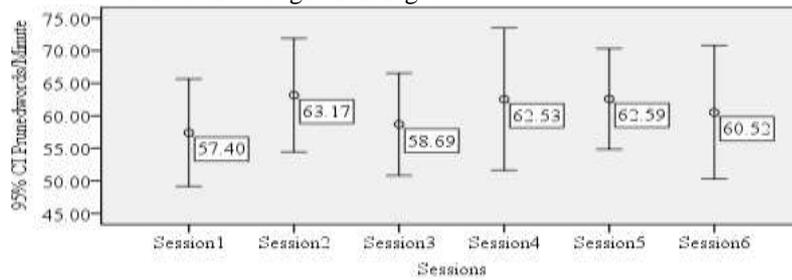

Figure 7
Extensive recast group's temporal fluency measurement

As it is depicted, there were unsteady rises and falls in temporal fluency of the extensive recast group between sessions 1 and 6.

Finally, the intensive recast group's spoken temporal fluency, pruned speech rate, between sessions 1 and 6 is given in Figure 8.
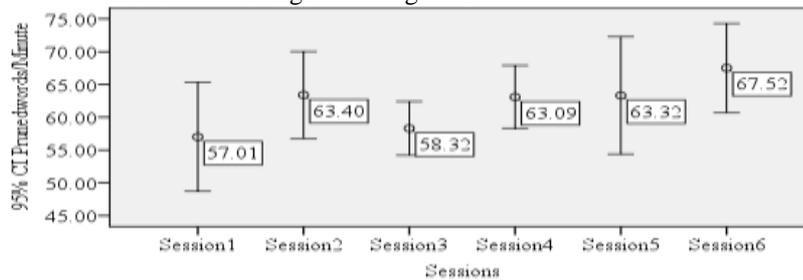

Figure 8
Intensive recast group's temporal fluency measurement

As it is depicted, there were unsteady changes in temporal fluency of the intensive recast group between sessions 1 and 3, but there was a steady rise in it between sessions 3 and 6.
To summarize, temporal fluency of the control group between sessions 1 and 4 declined. However, it increased between sessions 4 and 6. There was a steady rise in temporal fluency of the delayed explicit metalinguistic group between sessions 1 and 4, but it unsteadily changed between sessions 4 and 6. There were also unsteady rises and falls in temporal fluency of the extensive recast group between sessions 1 and 6. There were unsteady changes in temporal fluency of the intensive recast group between sessions 1 and 3, but there was a steady rise in it between sessions 3 and 6. Considering the fact that the comparisons were based on sessions 1 and 6, only these sessions' results are

comprehensively presented. Consequently, the descriptive statistics of all groups' temporal fluency scores, pruned speech rate, of sessions 1 and 6 is given in Table 2.

Table 2
Descriptive statistics of groups' spoken temporal fluency in sessions 1 and 6

Sessions		*n*	*Mean*	*SD*	*95% Confidence Interval*	
					Lower Bound	Upper Bound
Session 1	Control	16	69.20	19.34	58.91	79.52
	Delayed Explicit	17	55.30	19.72	45.17	65.44
	Extensive Recast	16	57.40	15.44	49.18	65.63
	Intensive Recast	15	57.01	14.98	48.73	65.31
Session 6	Control	18	69.92	21.62	59.17	80.67
	Delayed Explicit	18	62.08	16.96	53.66	70.52
	Extensive Recast	11	60.52	15.27	50.27	70.79
	Intensive Recast	14	67.52	11.77	60.73	74.32

To summarize the descriptive statistics of temporal measures of fluency of the spoken production, it should be restated that in all groups the temporal fluency of the spoken production in session 6 was bigger than that of session 1. In other words, the pruned speech rate of session 6 was bigger than that of session 1. In addition, the control group had the highest temporal fluency, pruned speech rate in session 6. Furthermore, the results of the one-way ANOVAs used to analyze the participants' spoken fluency revealed that there were insignificant differences among all groups in both sessions 1($F_{3, 60}$ = 2.13, p = .11) and 6 ($F_{3, 60}$ = .99, p = .41). In other words, the presence, absence, or type of CF had insignificant effect on the temporal fluency of the spoken production.

The Results of the Correlational Analyses of the Accuracy and Fluency Measurements in Sessions 1 and 6.

The results of the correlations of the general accuracy, error free clauses, and temporal fluency, pruned speech rate, of all groups in sessions 1 and 6 revealed that in all groups and in both sessions 1 and 6 the correlations were positive and insignificant (p > 0.05). In other words different CF conditions had insignificant effects on the correlations between spoken general accuracy and temporal fluency.

The Results of the Test of the Trade-off Hypothesis

In contrast, the results of the correlations between general accuracy, error free clauses, and temporal fluency, pruned speech rate, revealed that the correlation between general accuracy, error free clauses, and temporal fluency, pruned speech rate, was positive, and significant (r = .42) with a small effect size (R^2 = 0.18). In other words, as the number of error free clauses increased, the number of pruned words per min increased.

DISCUSSION

There were two purposes for the current study. The first one was to investigate the effects of the presence or absence, the type, the focus, and the timing of the CF on EFL learners' spoken general accuracy and temporal fluency. The other one was to study the presence or absence of the trade-off among the spoken accuracy and fluency and how it

would be affected by different CF conditions. Considering the first purpose, two research questions were asked.

The first research question was based on the effects of the presence or absence, type, focus, and timing of the CF on EFL learners' spoken general accuracy. The results of the descriptive statistics revealed that all groups were different and both the extensive recast and delayed explicit metalinguistic groups' spoken general accuracy in session 6 was higher than that of session 1 and the delayed explicit had the highest general accuracy in session 6. However, these differences weren't strong enough to reach a statistical significance. In other words, there were no significant differences among groups with respect to their production of the error free clauses.

The logical explanation for these results can be related to both the purpose of this research question and the nature of the accuracy measurement. On the one hand, the focus was on the simple past tense and, consequently, two of the participating groups named intensive recast and the delayed explicit recast groups, targeted it. On the other hand, the accuracy measurement was a general accuracy measurement and the only group which extensively covered all errors was the extensive recast group. Considering these two facts, the only group which had the possibility of developing the general accuracy was the extensive recast group. However, it couldn't develop the general accuracy and it had even a lower mean than the delayed explicit CF group. It can be explained with respect to the hypothesis that "discourse mode is likely to affect the particular linguistic forms a learner uses in performing a task" (Ellis, 2003, p.92). Consequently, a story-retelling task will lead more naturally to the use of the past tense.

This frequent use of past tense might have some consequences. The most important consequence, relevant to the current study, might be the frequency of the simple past tense errors. That is, due to the highest frequency of the simple past tense in the story-retelling tasks, they might potentially be the highest frequent errors. Indeed, even in the extensive recast group which was based on covering all errors, most of the errors might have been the simple past tense errors.

As a result, in the extensive recast group which was formed on the basis of correcting all errors, the errors might have mostly been the past tense errors and this type of CF had insignificant effects on spoken general accuracy. The results are in line with Kim and Mathes' (2001) study which revealed insignificant differences in the scores of the explicit and implicit groups. These results provide support for previous CF studies (e.g., Chehr Azad, Farrokhi, & Zohrabi, 2017; Doughty & Varela, 1998) which revealed that CF is beneficial to L2 learning when it specifically targets and measures particular forms. They are also in line with Rahimpour, Salimi, and farrokhi's (2012) study, investigating the effects of intensive and extensive focus on form strategies on female pre-intermediate EFL learners' oral accuracy, which revealed no differences between the performances of two groups in terms of accuracy in oral narrative task.

The second research question was based on the effects of the different CF conditions on the spoken temporal fluency of the EFL learners. To answer this question, it was needed to choose one of the temporal fluency indices. To this end, a correlational matrix was

created and its results revealed positive and significant correlations between temporal fluency indices and the effect size was very large. That is, the larger the pruned speech rate, the larger the unpruned speech rate.

The results of the descriptive statistics revealed that in all groups the temporal fluency of the spoken production in session 6 was bigger than that of session 1. In other words, the pruned speech rate of session 6 was bigger than that of session 1. In addition, the control group had the highest temporal fluency, pruned speech rate, in session 6. Considering the results of the inferential statistics, it was discovered that although the temporal fluency of all groups were different, the differences were not big enough to reach a statistical significance. In other words, the presence, absence, or type of CF had no significant effect on the temporal fluency of the spoken production. Despite the absence of significant differences, all groups having higher fluency in session 6 than in session 1 can be explained with respect to the nature of the task repetition, defined as "the repetition of the same or slightly altered task, whether the whole tasks, or parts of a task" (Bygate & Samuda, 2005, p.43). According to some researchers (Bygate, 2001; Ellis, 2003), task repetition can increase learners' fluency and complexity. The results are in line with Sato and Lyster's (2012) study which demonstrated that the presence or absence of the CF had no significant effect on the development of the learners' spoken fluency. They are also in line with Seyed Motahai and Ghasemi Nik Manesh's (2014) study which revealed that the type of the CF had no influence on impulsive and reflective EFL learners' spoken fluency. These results are in contrast with Rahimi and Vahid Dastjerdi's (2012) study which discovered that the CF type had a significant effect on the intermediate EFL learners' oral fluency and that the delayed CF was significantly more effective than the immediate CF for the development of their oral fluency.

The other purpose of the study was to examine the presence or absence of the trade-off among the spoken accuracy and fluency and how it would be affected by different CF conditions. Considering the nature of the trade-off, the results of the correlational analyses, based on all groups' spoken production in all sessions of the study, revealed a positive and significant correlation between general accuracy, error free clauses, and temporal fluency, pruned speech rate. These results are in contrast with the trade-off hypothesis that "committing attention to one area, other things being equal, might cause lower performance in others" (Skehan, 1998, p. 112). It is also in contrast with Skehan's (1998) suggestions that there might be a tension between form, complexity and accuracy, on the one hand, and fluency, on the other hand. The results are inconsistent with previous studies (Ahmadian & Tavakoli, 2011; Michel, Kuiken, & Vedder, 2007; Wendel, 1997; Yuan & Ellis, 2003) which revealed a trade-off between accuracy and fluency. With respect to the effects of different CF conditions on the trade-off between AF, the results of the correlational analyses revealed that different CF conditions had no significant effects on the positive and significant correlations between spoken general accuracy and temporal fluency. This finding indicating that different CF conditions have no effect on the relationship between accuracy and fluency, is consistent with Chehr Azad et al.'s (2017) study which revealed insignificant effects of the CF conditions on the relationships between specific accuracy and repair fluency.

CONCLUSION

One of the limitations of the study was related to its length. That is, it was very short and lasted six sessions. Another limitation was related to the target of the study which was the simple past tense. Furthermore, CF types were input providing. In addition, only one task, story-retelling task, was used. The other limitation was related to the proficiency level of the participants. Consequently, further research, focusing on long treatments, other grammatical structures, output prompting CF types, different tasks, and different proficiency level participants, needs to be done to address these limitations. With respect to the pedagogical implications, it can be suggested that the provision of the different CF types, whether intensive or extensive, has no significant effects on the EFL pre-intermediate learners' spoken general accuracy and temporal fluency. In addition, repeated performance of story-retelling task is likely to develop the EFL pre-intermediate learners' spoken temporal fluency. There is no trade-off between spoken general accuracy and temporal fluency.

Based on the findings, it can be concluded that different CF conditions, particularly whether they are intensive or extensive type, have no significant effects on the EFL pre-intermediate level learners' spoken general accuracy in the story-retelling tasks. In addition, different CF conditions have insignificant effect on the EFL pre-intermediate learners' temporal fluency. However, repeated performance of story retelling tasks, irrespective of the presence or absence of CF, might have some insignificant but positive influences on the EFL learners' spoken temporal fluency. Considering the relationships between spoken general accuracy and temporal fluency, it should be suggested they are positively and significantly correlated. That is, the higher the spoken general accuracy, the higher the temporal fluency and vice versa. In addition, CF has no significant effect on their relationships.

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. doi: 10.1177/1362168810383329

Ammar , A., & Spada , N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, *28*, 543–574.

Ansarian, A. A., & Chehr Azad, M. H. (2015). Differential effects of focused and unfocused recasts on the EFL learners' oral accuracy. *Colomb. Appl. Linguist. J., 17*(1), 86-97.

Brumfit, G.J. (1984). *General English syllabus design*. Oxford: Pergamon Press.

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 a second language* (pp. 37–74). Philadelphia, PA: Benjamins.

Chehr Azad, M. H., Farrokhi, F., & Zohrabi, M. (2017). *The Impact of corrective feedback (CF) on Iranian EFL learners' spoken complexity, accuracy, and fluency (CAF) and trade-off among them.* Manuscript submitted for publication.

Doughty, C. J. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). New York: Cambridge University Press.

Dougthy, C, J., & Varela, E. (1998). Communicative focus on form. In C. J. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York: Cambridge University Press.

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305 – 352.

Ellis, R. (1991). *Second language acquisition and language pedagogy.* Clevedon: Multilingual Matters.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79–114). San Diego, CA: Academic Press.

Ellis, R. (2001). Investigating form-focused instruction. *Language Learning, 51*(1), 1–46.

Ellis, R. (2003). *Task based language learning and teaching*. Oxford, UK: Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (2009 a). Corrective feedback and teacher development. *L2 Journal*, *1*(1), 3–18.

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

Ellis. R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28,* 339–358.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36*, 353–371.

Farrokhi, F., & Chehr Azad, M. (2012). The effects of planned focus on form on Iranian EFL learners' oral accuracy. *World Journal of Education, 2*(1), 70-81.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 293–323.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition, 16*, 283–302.

Hill, L. A. (1988). *Steps to Understanding.* Oxford: Oxford University Press.

Hoseini Fatemi, A., & Harati, N., A. (2014). The Impact of Recast versus Prompts on the Grammatical Accuracy of Iranian EFL Learners' Speech. *Theory and Practice in Language Studies,4,* 532–543.

Kamiya, N. (2015). The effectiveness of intensive and extensive recasts on L2 acquisition for implicit and explicit knowledge. *Linguistics and Education, 29,* 59–72.

Kim, H., & Mathes, G. (2001). Explicit vs. implicit corrective feedback. *Korea TESOL Journal, 4,* 1-15.

Li, S. (2010).The Effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309–365. doi:10.1111/j.1467-9922.2010.00561.x

Loewen, S. (2011). Focus on form. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 576–592). London: Routledge.

Long. M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. Ginsberg, & C. K. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

Lyster, R. (2004).Differential effects of prompts and recasts in form- focused instruction. *Studies in Second Language Acquisition*, *26*, 399–432.

Lyster, R., & Ranta, L. (1997). Corrective feedback and the learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19,* 37–66.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Maftoon, P., & Kolahi, S. (2009). The impact of recasts on the syntactic accuracy of Iranian EFL university students' oral discourse. *The Journal of Applied Linguistics, 2*(2), 160–178.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching (IRAL), 45*, 241–259.

Nicholas, H., Lightbown, P., & Spada, N. (2001). *Recasts as feedback to language learners*. *Language Learning, 51*, 719–758.

Rahimi, A., & Vahid Dastjerdi, H. (2012). Impact of Immediate and Delayed Error Correction on EFL Learners' Oral Production: CAF. *Mediterranean Journal of Social Sciences, 3*(1), 45–54.

Rahimpour, M., Salimi, A., & Farrokhi, F. (2012). The effect of planned vs. unplanned form-focused strategies on L2 learners' accuracy in oral task performance. *Education Research Journal, 2,* 247–252.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 131–164). Amsterdam: Benjamins.

Saeedi, P. (2015). Unguided strategic planning, task structure, and L2 oral performance: Focusing on complexity, accuracy, and fluency. *Journal of Applied Linguistics and Language Research, 2*, 263–274.

Salimi, A. (2015).The Effect of Focus on Form and Task Complexity on L2 Learners' Oral Task Performance. *Advances in Language and Literary Studies, 6*(6), 54–62.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: monitoring, practice and proceduralization. *Studies in Second Language Acquisition*, *34*, 591–626.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied linguistics, 11,* 129–158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.

Seyed Motahari, M., & Ghasemi Nik Manesh, A. (2014).The comparative effect of explicit corrective feedback and clarification request feedback on impulsive and reflective EFL learners' oral fluency. *International Journal of Enhanced Research in Educational Development (IJERED), 2* (2), 32–42.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301–322). Oxford: Oxford University Press.

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37,* 556–569.

Skehan, P. (1996). A framework for the implication of task based instruction. *Applied Linguistics, 17,* 38–61.

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*, 510–532.doi:10.1093/applin/amp047

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*, 185–211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing condition on narrative retelling. *Language Learning*, *49*, 93–120.

Skehan, P., & Foster, P. (2001). Cognition and Tasks. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 183–205). Cambridge: Cambridge University Press.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In: R. Ellis (Ed.), *Planning and task-performance in a second language* (pp. 193–216). Amsterdam: Benjamins.

Vercellotti, M. L. (2012).*Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other* (Unpublished doctoral dissertation). University of Pittsburgh, USA.

Wendel, J. (1997). *Planning and second language narrative production* (Unpublished doctoral dissertation). Temple University, Japan.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24,* 1–27. doi: 10.1093/applin/24.1.1