

The Role of Direct Metalinguistic and Indirect Feedback Timing in Enhancing Speaking Accuracy Task Performance of Iranian EFL Learners

Elham Dehdari¹, Nasser Ghafoori^{2*}, Saeideh Ahangari³

¹Ph.D. Candidate, Department of English Language, Sarab Branch, Islamic Azad University, Sarab, Iran

^{2*}Assistant Professor, Department of English Language, Tabriz Branch, Islamic Azad University, Tabriz, Iran

³Associate Professor, Department of English Language, Tabriz Branch, Islamic Azad University, Tabriz, Iran

Received: October 30, 2023

Accepted: December 02, 2023

Abstract

The primary objective of this study was to investigate the impact of feedback timing on the improvement of Iranian EFL learners' speaking accuracy in direct metalinguistic and indirect feedback scenarios. By utilizing a quasi-experimental design, the investigator enlisted a total of 120 participants, consisting of 63 males and 57 females, with an average age of 21.5 years. These participants were then divided into four distinct groups: a direct-immediate feedback group comprising 30 students, a direct-delayed feedback group consisting of 30 students, an indirect-immediate feedback group comprising 30 students, and an indirect-delayed feedback group consisting of 30 students. The data collection process involved the utilization of three distinct instruments: the Oxford Placement Test (OPT), a speaking accuracy test, and a series of two-way exchange tasks. Following the homogenization of participants based on the OPT test, a pretest was administered to assess their speaking accuracy. In the treatment phase, the group receiving indirect instant feedback promptly rectified errors by reformulating them into the proper form upon occurrence. In the experimental condition of indirect delayed feedback, the instructor meticulously documented the errors made by each student along with their respective names, to deliver personalized feedback to each learner. In the experimental group focused on direct metalinguistic immediate feedback, participants were presented with explanations and examples illustrating the right form of the errors they had made. In the metalinguistic delayed group, the instructor would record the nature of errors and the names of the pupils to offer explanations and provide accurate examples tailored to each student. During the post-test phase, the participants underwent a re-administration of the speaking accuracy test, and the resulting scores were duly recorded. The data that was gathered was subjected to both descriptive and inferential analysis using Two-way ANOVA. The findings from the data analysis indicate a notable interaction between the timing of feedback and the type of input about their impact on the speaking accuracy of Iranian intermediate English as a Foreign Language (EFL) learners.

Keywords: Feedback, meta-linguist feedback, indirect feedback, speaking accuracy

INTRODUCTION

Feedback is an often-seen occurrence in language training programs. The influence of corrective feedback (CF) types on language

learning has been recognized as a significant aspect in the field of language feedback research (Park, 2010). The authors Macintyre, Burns, and Jessome (2011) examine the effects of various corrective feedback (CF) strategies on students' willingness to communicate,

*Corresponding Author's

Email: ghafoori@iaut.ac.ir

motivation, and anxiety. They explore how CF can either enhance or diminish these factors, depending on the expectations around its use and how it is provided.

Intercultural communication stands as a prominent objective in the pursuit of acquiring additional languages. According to Scrivener (2011), Chastain (1971) places significant emphasis on the development of speaking skills as the most crucial competency to be attained when acquiring proficiency in a foreign language. According to his assertion, a high degree of proficiency in oral communication helps ensure the development of proficiency in other areas such as reading, writing, and listening. According to the author, it is argued that oral communication has two primary functions in the context of English as a Foreign Language (EFL) learning. These functions include reinforcing the acquisition and utilization of the learners' existing language ability.

According to Skehan (2009), the development of speaking skills, as well as writing abilities, is intricately connected to task performance. Previous research has substantiated the notion that engaging in tasks contributes to the acquisition and enhancement of oral communication abilities (e.g., Robinson, 1995). Unfortunately, as a result of certain teachers' insufficient understanding of the appropriate utilization of corrective feedback (CF), many learners experience a decline in their self-assurance in oral communication and subsequently encounter academic difficulties. The cause of this failure stems from the fact that feedback is not a singular concept; rather, it can be delivered through various means, such as direct and indirect input. Prior research (e.g., Karimi & Asadnia, 2015) has established that there exists a notable difference in the utilization of different types of corrective feedback (CF) by teachers, depending on the source of the input (i.e., teacher, self, peer). The inclusion of feedback-related variables, such as timing, feedback source, and type, has the potential to yield varying outcomes. While it is important to acknowledge the significance of analytic studies, a comprehensive response to the inquiry regarding the efficiency of corrective feedback necessitates the consideration of a broader range of contextual and learner

characteristics within the taught second language acquisition (SLA) setting (Ellis, 2010). The existing body of research lacks clarity regarding whether reactive attention to linguistic form should be employed during the execution of a task or if it can be deferred until the activity has been concluded. Limited discourse on this topic can be observed within the realm of Task-Based Language Teaching communities, as noted by Ellis (2017). Additionally, the temporal aspect of providing feedback is a significant concern, as it is commonly advised in pedagogy that feedback should be delivered promptly for accuracy-focused tasks, but delayed for tasks focused on fluency (Hedge, 2000; Scrivener, 2011). Hence, given the importance of feedback and the existing knowledge gaps surrounding it, this study aims to examine the efficacy of different feedback timing (immediate and delayed) about the type of feedback (direct metalinguistic and indirect feedback).

To achieve this objective, the following research questions were formulated:

RQ1: *Does feedback timing have any significant effect on intermediate EFL learners speaking accuracy in task performance?*

RQ2: *Does feedback type have any significant effect on intermediate EFL learners' speaking accuracy in task performance?*

RQ3: *Is there any interaction between feedback timing and feedback type on intermediate EFL learners' speaking in task performance?*

Review of Literature

The topic of feedback timing has historically received limited attention in SLA research literature, with a renewed interest emerging in recent times (Ellis, 2017). In the pre-communicative language teaching era, scholars frequently deliberated on the optimal timing for providing feedback, with certain theorists advocating for rapid feedback while others proposed a delayed approach. Nevertheless, subsequent to that period, the discourse surrounding the timing of feedback has significantly waned. The perception of when feedback should be provided has been significantly shaped by shifts in the understanding of language acquisition processes. During the 1950s and 1960s, scholars

advocated for the strict avoidance of errors, likening them to sins. They further emphasized the importance of promptly addressing any errors that did occur (Hendrickson, 1978). In accordance with the principles of behaviorism, scholars such as Brooks (1960) advocated for minimizing the temporal gap between learners' errors and teachers' provision of a correct model. Theorists expressed concerns regarding the potential for improper linguistic behavior to become deeply rooted if learners were exposed to it for an extended period without any intervention.

A change in perspective emerged when Corder (1967) made a significant acknowledgment of the significance of errors as manifestations of the progress learners were achieving in their second language acquisition. Subsequently, educators were encouraged to engage in the analysis of errors with the aim of discerning the hypotheses that learners may be testing regarding the second language (L2) through the manifestation of said errors. According to Fanselow (1977), new teachers may need to postpone providing feedback for a period of up to one day in order to thoroughly assess the nature of an issue and determine the most effective approach for addressing it. Similarly, Chastain (1971) proposed the idea that teachers could engage in the practice of reviewing typical errors subsequent to the completion of communicative activities. In contrast, Allwright (1975) stated that teachers should refrain from waiting and instead engage in the prompt analysis and provision of immediate feedback.

The presence of divergent perspectives on the timing of feedback can be attributed to various factors, one of which is the sort of feedback provided. Several research, such as the one conducted by Rosa and Leow (2004), have demonstrated that explicit feedback has a greater impact compared to implicit feedback. Similarly, the study conducted by Ellis, Loewen, and Erlam (2006) revealed that the impact of explicit feedback surpassed that of implicit input. However, some investigations (e.g., Fu & Nassaji, 2016) have reported no significant differences. Only one study (Leeman, 2003) has reported that implicit corrective feedback had a greater impact compared to explicit input.

Once again, this contradictory finding may be ascribed to the temporal aspect of feedback. Indeed, the aforementioned research did not consider the influence of feedback timing on language-related abilities, such as the accuracy of speaking. Numerous recent research have been conducted to examine the effects of quick feedback compared to delayed feedback on the process of language acquisition. As an illustration, a study conducted by Li (2015) revealed that providing rapid feedback yielded superior results in enhancing grammar accuracy compared to delayed input.

According to Wang's (2017) study, it was determined that quick feedback yielded more efficacy in enhancing speaking fluency compared to delayed input. Nevertheless, alternative research has indicated that there is no substantial difference observed between immediate and delayed input. An empirical investigation conducted by Zhang (2018) revealed that both immediate and delayed feedback had comparable efficacy in enhancing vocabulary acquisition. Chen (2019) conducted a study that revealed that both immediate and delayed feedback demonstrated comparable efficacy in enhancing writing accuracy. The efficiency of feedback time can be influenced by various aspects, such as the nature of the feedback provided, the competence level of the learner, and the learner's unique learning style. For instance, certain forms of feedback, such as explicit feedback, may exhibit greater efficacy when delivered promptly, but other forms of feedback, such as implicit feedback, may demonstrate enhanced effectiveness when administered gradually. Furthermore, individuals who possess advanced levels of skill may potentially derive greater advantages from delayed input compared to those with lesser levels of proficiency. Ultimately, individuals with varying learning styles may exhibit a preference for distinct forms of feedback timing.

According to Lindsay and Knight (2006), those aspiring to become proficient speakers must consider several factors. These include the ability to generate coherent speech, effectively express ideas, articulate information, and adapt speech to various contexts and situations with both accuracy and fluency. Accuracy is a

linguistic proficiency that pertains to the capacity to produce grammatically accurate phrases. It encompasses the appropriate utilization of grammar, vocabulary, and other language skills. In order to achieve accuracy, the learner should focus on the proper form of expression, namely, ensuring accuracy. The simultaneous focus on form and meaning often poses a challenge for students in maintaining concentration. Achieving accuracy necessitates careful attention and the allocation of sufficient time. The proposition put forth by researchers posits that pupils' accuracy levels are positively correlated with the amount of time they have available. Language learners often encounter challenges when it comes to pronunciation, as words that are difficult to speak tend to be more challenging to learn. Words that may pose challenges for certain groups of students are ones that contain sounds unfamiliar to them (Thornbury, 2000). According to Long (1996), negative evidence is a prominent occurrence in the context of feedback. Feedback is given to learners through negative evidence, which assesses the clarity, comprehensibility, and accuracy of their production (Robinson, 1995). The ability to produce speech that is both understandable and accurate is of great importance, as it serves as a crucial factor in preventing communicative errors that hinder successful engagement with individuals.

A growing body of research has been dedicated to examining the potential differential effectiveness of various types of corrective feedback (CF) in enhancing the writing accuracy of second language (L2) learners. Upon examining several studies, Truscott (1999) observed that none of them yielded statistically significant variations among the various treatment groups, including those receiving content comments exclusively, error correction exclusively, a combination of content comments and error correction, or error identification without correction. However, it is important to exercise caution when drawing conclusions from these findings, as additional research exploring alternative feedback distinctions suggests that the aforementioned conclusion may not be entirely reliable at this point. Numerous research have undertaken the task of differentiating between

direct and indirect feedback systems, with the aim of examining their efficacy in enhancing accuracy (Robb, Ross, & Shortreed, 1986). Direct or explicit feedback is characterized by the teacher's identification of an error and subsequent provision of the correct form. On the other hand, indirect tactics involve the teacher indicating that an error has been made without providing a correction, so requiring the student to independently identify and rectify the mistake.

Furthermore, scholarly investigations investigating the impact of indirect feedback strategies have commonly made an additional differentiation between approaches that employ or do not employ a code. Coded feedback is a method that precisely identifies the specific position of an error, while also providing a code to indicate the type of error committed. For instance, the code "PS" signifies an error related to the use or form of the past simple tense. The term "uncoded feedback" pertains to situations in which an instructor highlights an error, encircles an error, or keeps a count of errors in the margin while leaving it up to the student to identify and rectify the problem in question. While some prior investigations in this field lacked a control group, the research conducted by Ferris and Roberts (2001) featured control groups that were not exposed to corrective feedback (CF). The research conducted by Lee on English as a Foreign Language (EFL) college students in Hong Kong revealed a noteworthy impact observed in the group whose faults were underlined, as opposed to the groups that did not receive any corrective feedback (CF) or just received a little check. In their study, Ferris and Roberts (2001) investigated the impact of three distinct feedback treatments on performance outcomes. These treatments included marking errors with codes, underlining errors without additional marking or labeling, and providing no error feedback. The results revealed that both error feedback groups demonstrated significantly higher performance levels compared to the control group that received no feedback. However, similar to the findings of Robb et al. (1986), no significant differences were observed between the group receiving coded feedback and the group not receiving coded feedback. Moreover, it is important to

acknowledge that the study conducted by Ferris and Roberts (2001) focused on the examination of text changes as opposed to the creation of new written compositions over a period of time.

Research in the field of feedback has also uncovered variations in students' responses to feedback provided by their peers compared to feedback provided by adults. The study conducted by Praver, Rouault, and Eidswick (2011) examined the impact of gender on individuals' responses to feedback. However, it is important to note that this relationship is influenced by age. The activities for providing feedback can exhibit variations across several dimensions, encompassing diverse curriculum areas or courses. Various forms of feedback may be deemed suitable throughout the language acquisition process.

METHODOLOGY

Participants

The research was carried out over the summer of 2019 at the Iranmehr language institutes located in Tehran, Iran. The participants of this study consisted of intermediate English as a Foreign Language (EFL) learners. The classes were conducted over a period of 16 sessions, occurring twice every week. Utilizing a quasi-experimental methodology, the investigator enlisted a sample of 120 individuals from a larger group of 200 students subsequent to administering an Oxford Placement Test (OPT). Hence, a total of 120 students were chosen to participate in the study, divided into four distinct experimental groups. The first group consisted of 30 students who received direct-immediate feedback, while the second group comprised 30 students who received direct-delayed feedback. Similarly, the third group consisted of 30 students who received indirect-immediate feedback, and the fourth group included 30 students who received indirect-delayed feedback. The study included a total of 120 individuals, consisting of 63 males and 57 females. The age range of the participants in the study spanned from 15 to 31 years, with a mean age of 21.5 years. Three participants were excluded from the intervention process. The selection of a single teacher for the instruction of all students was made. The reason for selecting the same

teacher for all pupils was to mitigate any potential confounding variables that could arise from teachers' personal characteristics.

Instruments

Oxford Placement Test (OPT)

In order to assess potential disparities among the study participants, the researchers administered the OPT test (version I), an English language assessment developed by Oxford University Press and the University of Cambridge Local Examinations Syndicate, to the students. The examination serves as an assessment that showcases one's proficiency in utilizing the English language for various everyday communication functions. The instrument utilized in this investigation is of primary importance. The assessment has a total of 60 items that pertain to language structures, presented in a multiple-choice format. As per the test makers' guidelines, those who obtain scores ranging from 37 to 46 are classified as intermediate. The participants were given a time limit of 30 minutes to respond to the questions.

Speaking Accuracy Test

To examine the impact of immediate and delayed feedback on the speaking accuracy of participants, the researcher utilized a speaking accuracy test derived from the PET test (2018). To achieve this objective, the participants were instructed to engage in a collaborative task with a partner under a time limit of four minutes. The vocalizations were captured for subsequent examination. The focus of this study was the utilization of English tense by students. The researchers calculated the ratio of correct verbs to the total number of verbs used in their speech. The ratio in question was considered to be the test score for the accuracy of the participants (Yuan & Ellis, 2003). Previous scholars (Assasi, 2018) have evaluated the reliability, validity, and usability of the test within the Iranian setting. To determine the dependability, two raters were enlisted to evaluate the speaking accuracy exams. Hence, the assessment of dependability was conducted by examining interrater reliability. The measure's validity was further substantiated through a panel debate.

A group of TEFL professionals was solicited to provide their perspectives on the test's validity.

Language Tasks

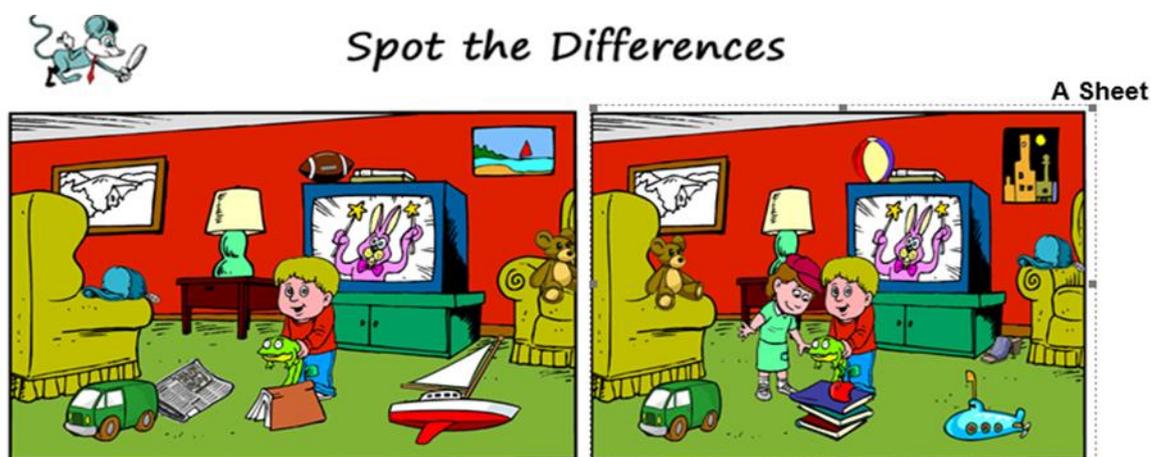
During the treatment participants performed one task each session. A two-way exchange

task (Spot-the-differences) that elicits collaborative interaction was chosen for the current study. Some examples of the topics which were given to the participants are provided in Table 1:

Table 1
Topic, task type, and the description of the task

Topic	Task Type	Description of the task
Persian vs. American culture	Spot-the-differences	Goal-oriented; convergent; one closed outcome
Male and females	Spot-the-differences	Goal-oriented; convergent; one closed outcome

*Description of each task is based on Pica, Kanagy, and Falodum (1993)



Describe your picture to your partner and try to find as many differences as you can. When you are finished compare your pictures and write down the differences that you found.

These two-way exchanges promote negotiation of meaning and form (Skehan, 2003). For instance, one of the goal-oriented activities is for the participants to work together to identify 15 differences between two drawings of a messy room. Specific lexical items or grammar points would likely be required to achieve mutual comprehension.

Procedure

The primary objective of the current study was to examine the impact of a specific variable on other variables. To achieve this objective, a quasi-experimental research design with pre-test-posttest measures was employed. Following the selection of the study's final participants

using the OPT test, a speaking accuracy test was administered as a pretest. This test had the participants doing a task. To achieve this objective, the participants were instructed to engage in a collaborative task with a partner within a time frame of four minutes. The vocalizations were captured to do subsequent analysis. The focus of this study was on the utilization of English tense by students. To determine the accuracy of their verb usage, the researchers divided the number of correct verbs by the total number of verbs present in their speech (Yuan & Ellis, 2003). The responses were documented in order to be subjected to analysis to assess their accuracy. Subsequently, participants were allocated to one of the four groups

according to the specific feedback they were to get (either direct or indirect) and the timing of the input (either immediate or delayed). In the group that received indirect rapid feedback, participants were instructed to promptly repair any errors they made by reformulating them into the correct form. As previously stated, the faults within this group were rectified during the execution of the task. In the experimental condition of indirect delayed feedback, the instructor would record instances of errors made by students, along with their respective names, to subsequently deliver feedback tailored to each individual student. According to Li, Zhu, and Ellis (2016), the feedback was deemed delayed as it was given after the completion of the assignment. Following the conclusion of the peer interaction, the teacher would proceed to engage in a discussion regarding the inaccuracies. As previously stated, the faults within this group were rectified after the completion of the work. The participants in the direct metalinguistic immediate feedback group were given explanations and examples of the right form of errors they made (Nassaji, 2015). As previously stated, the faults within this group were rectified during the execution of the task.

In the final group, referred to as the direct metalinguistic delayed group, the instructor would record the nature of errors made by students, along with their respective names, to offer explanations and provide accurate examples tailored to each student. As previously indicated, the feedback was deemed delayed as it was administered after the completion of the activity (Li, Zhu, & Ellis, 2016). Following the conclusion of the peer engagement, the teacher

would proceed to elucidate the faults. Following the completion of the treatment, the subjects underwent a posttest in the form of a speaking accuracy test. It is necessary to include that the emphasis of the accuracy component in speaking task performance was placed on the utilization of English tenses. To initiate the analysis of the results, the researcher employed the Smirnov-Kolmogorov test to assess the homogeneity of the scores across the four groups. Following the attainment of good outcomes, the researcher proceeded to conduct a comparative analysis of the means of the test findings. Subsequently, the acquired data was subjected to an examination including both descriptive and inferential statistical techniques. A Two-way ANOVA test was used to examine the disparities between the groups in terms of pre and post-test results.

RESULTS

To measure the effects and the interaction of the two independent variables on participants' speaking accuracy, there was a need to compare pre-test and post-test speaking accuracy scores to see if any change in their scores had occurred or not.

Descriptive statistics, as shown in Table 2, were run to notice if there was any speaking accuracy difference among the four groups, before the treatment. Table 1 provides the mean and standard deviation for each combination of the groups of the independent variables. In addition, the table provides "Total" rows, which allows means and standard deviations for groups only split by one independent variable.

Table 2
Descriptive statistics related to pretest speaking accuracy scores of the four groups

Descriptive Statistics				
Dependent Variable: pre-accuracy				
Feedback type	Feedback timing	Mean	Std. Deviation	N
indirect	immediate	.59	.13	30
	delayed	.58	.11	30
	Total	.58	.12	60
direct	immediate	.58	.10	30
	delayed	.57	.11	30
	Total	.57	.10	60
Total	immediate	.58	.11	60
	delayed	.58	.11	60
	Total	.58	.11	120

According to the data shown in Table 2, it can be observed that the initial speaking accuracy scores of the four groups were approximately equal. Based on the data shown in Table 1, it can be observed that the group members in the indirect immediate condition had the best scores in terms of speaking accuracy ($M = 0.592$). Conversely, the participants in the direct delayed condition obtained the lowest results ($M = 0.576$). While the mean scores of speaking accuracy varied among the four groups, it is necessary to assess the statistical significance of these variances. Certainly, it is possible that the disparities in scores may

have been inconsequential. The utilization of Two-way ANOVA was employed in order to compare the outcomes, given the presence of two factors each with two levels. Prior to conducting a comparison of pretest and posttest data using Two-way ANOVA, it was imperative to ensure that the six assumptions of Two-way ANOVA were satisfied. The pretest scores of the four groups were compared using the Two-Way ANOVA approach, given that the six assumptions had been satisfied. Table 3 displays the outcomes of the Two-way Analysis of Variance (ANOVA) examination.

Table 3
Two-way ANOVA results

Tests of Between-Subjects Effects						
Dependent Variable: speaking accuracy						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.004 ^a	3	.00	.094	.96	.00
Intercept	40.84	1	40.84	2.97	.00	.96
Feedback type	.00	1	.00	.17	.68	.00
Feedback timing	.001	1	.00	.10	.75	.00
Feedback type * Feedback timing	.000	1	.00	.01	.91	.00

a. R Squared = .002 (Adjusted R Squared = -.023)

The particular rows to be considered are the "Feedback type pre", "Feedback timing pre" and "Feedback type pre * Feedback timing pre" rows. These rows inform us whether the independent variables (the "feedback type" and "feedback timing" rows) and their interaction (the "feedback type*feedback timing" row) do not have a statistically significant difference in mean accuracy scores. Under the "Sig." column, it can be observed that there is no statistically significant interaction between feedback timing and type at the $p = .91 > 0.01$ level. This means that there is no statistically significant

difference between the groups. Simple main effects analysis also showed that neither feedback timing nor feedback type was significantly different (feedback type: $p = .68 > 0.01$; feedback timing: $p = .75 > 0.01$). To sum it up, it can be inferred from the table above that there was no statistically significant difference in mean accuracy score between all four groups.

To better understand the similarity of speaking accuracy scores in the pretest, the plot of the mean "speaking accuracy" score for each combination of groups of "feedback type" and "feedback timing" is plotted below as Figure 1.

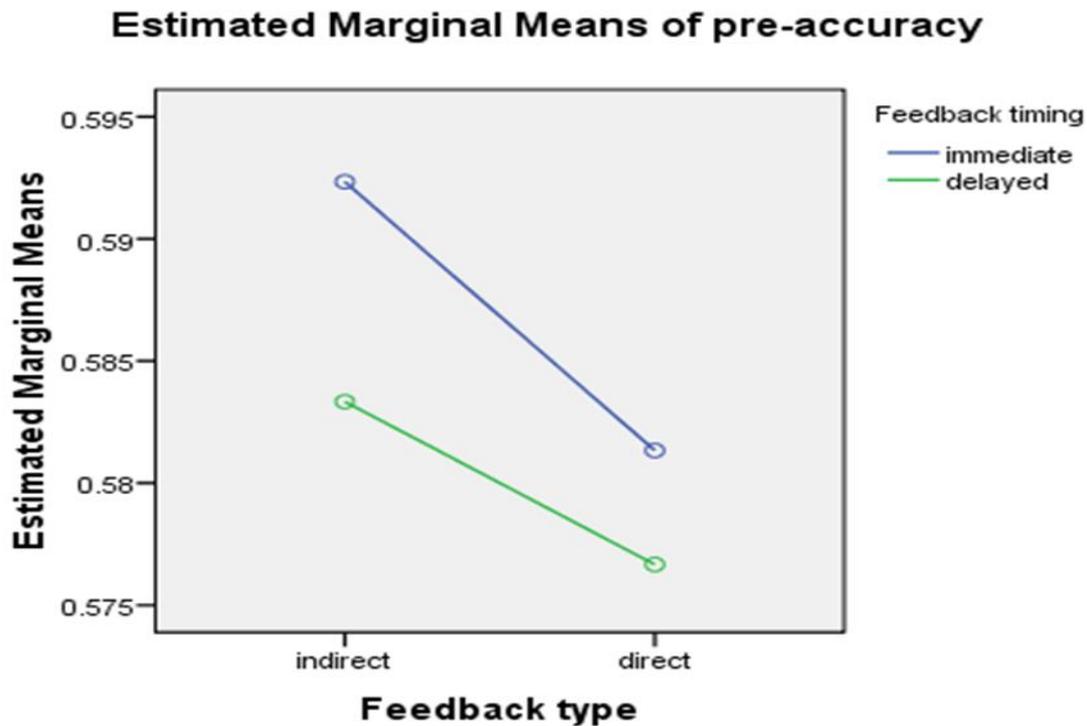


Figure 1
Plot of pretest results

As is commonly noticed in the graphical representations, the presence of an interaction effect is typically shown by the occurrence of non-parallel lines. The graph illustrates a notable observation wherein the lines exhibit a parallel relationship, lacking any instances of intersection. Thus far, it has been ascertained that there existed no notable difference among the four groups prior to the commencement of the treatment. To assess the impact of treatment on speaking accuracy, it was necessary to conduct a statistical analysis to determine the significance of the difference in post-test accuracy ratings across the four groups. To assess the statistical difference, a Two-way ANOVA should have been conducted due to the presence of two independent variables and one dependent variable. In a manner akin to the approach taken for the pretest scores, it was necessary to ascertain whether there were any discernible differences in the mean accuracy scores among the four groups. Prior to doing a Two-way ANOVA analysis to assess the statistical significance of differences, it is imperative to ascertain the presence or absence of any variations. To achieve this objective, descriptive statistics

were utilized to examine whether there were any variations in speaking accuracy among the four groups following the treatment. These statistics are presented in Table 4.

Table 4 presents the average and standard deviation values for each grouping of the independent variables. Furthermore, the table includes "Total" rows, which enable the determination of averages and standard deviations for groups that are divided by only one independent variable or not divided at all. According to the data presented in Table 4, there were notable discrepancies in the speaking accuracy scores across the four groups following the treatment. Based on the data presented in Table 4, it can be observed that the group members in the indirect delayed condition achieved the best results in terms of speaking accuracy, with a mean score of 0.7. Conversely, the participants in the direct instantaneous condition had the lowest scores, with a mean score of 0.59. While the mean scores of speaking accuracy varied among the four groups, it is necessary to assess the statistical significance of these variances. The disparities in scores may have been inconsequential.

Table 4
Descriptive statistics related to post-test speaking accuracy scores of the four groups

Descriptive Statistics				
Dependent Variable: post-accuracy				
Feedback type	Feedback timing	Mean	Std. Deviation	N
Indirect	Immediate	.60	.12	30
	delayed	.70	.07	30
	Total	.65	.11	60
Direct	immediate	.59	.12	30
	delayed	.59	.11	30
	Total	.59	.11	60
Total	immediate	.60	.12	60
	delayed	.65	.10	60
	Total	.62	.11	120

To check the statistical significance of the score differences, Two-way ANOVA needed to be run. As maintained before, the six assumptions of Two-way ANOVA were checked. Since the assumptions of the Two-way

ANOVA were met and the mean scores of posttest accuracy across the four groups were different, Two-way ANOVA was run and the results related to it are presented in Table 5.

Table 5
Two-way ANOVA results

Tests of Between-Subjects Effects						
Dependent Variable: speaking accuracy						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.23 ^a	3	.07	6.36	.00	.14
Intercept	46.93	1	46.93	3.78	.00	.97
Feedback type	.08	1	.08	6.91	.01	.05
Feedback timing	.07	1	.07	6.24	.01	.05
Feedback type* feedback timing	.07	1	.07	5.92	.01	.04

a. R Squared = .141 (Adjusted R Squared = .119)

The rows that were mandated for reporting from this table included the "Feedback type post," "Feedback timing post," and "Feedback type post*Feedback timing post" fields. The purpose of these rows is to determine if there is a statistically significant difference in mean accuracy scores for the independent variables, namely the "feedback type" and "feedback timing" rows, as well as their interaction, represented by the "feedback type*feedback timing" row. The data in the "Sig." column indicates a statistically significant distinction at the $p = .01$ level concerning the feedback kind. This finding indicates a notable difference in the post-test speaking accuracy scores between the indirect and direct groups. Furthermore, there

exists a statistically significant impact at the significance level of $p = .01$ for the timing of feedback. This finding indicates a notable difference in the post-test speaking accuracy scores between the groups that received quick feedback and those that received delayed input. Ultimately, and of greater significance, it may be inferred that there existed a statistically significant interaction between the impacts of feedback type and feedback time on the accuracy of speaking, with a p -value of .01. In conclusion, the data shown in the table indicates a statistically significant variation in the average accuracy scores among the four groups. The effect size and significance of the difference are presented in Table 5.

To determine the extent of variance in mean accuracy scores attributable to the independent variable, it is necessary to report the effect size, specifically the partial eta squared value found in the last column of the aforementioned table. The feedback type has been observed to have an effect size of 0.56. This indicates that the type of feedback accounts for 0.056 of the difference seen among the four groups. Furthermore, previous research has indicated that a modest effect size of 0.051 is associated with the proportion of score variance that can be attributed to differences in feedback time. In conclusion, the impact size of 0.049 suggests that approximately 4.9% of

the variability in accuracy scores can be attributed to the combined influence of feedback type and feedback time. It is important to note that all four effect sizes are classified as small. Nevertheless, it is important to note that eta square provides an estimation of the effect size solely inside the sample, rather than in the entire population.

To enhance comprehension of the difference in speaking accuracy scores between the post-test and the interaction involving feedback type and feedback timing, Figure 2 depicts the average scores for speaking accuracy across various combinations of feedback type and feedback time groups.

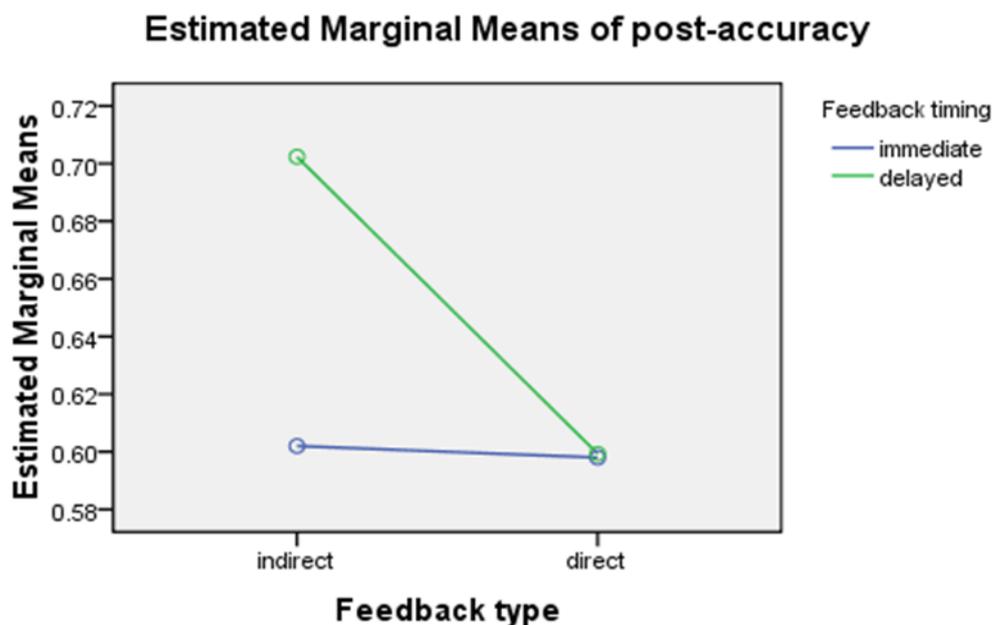


Figure 2
Plot of post-test accuracy results

Typically, an interaction effect manifests as a collection of non-parallel lines. The graph visually indicates that the lines intersect at a specific place. In essence, the graph illustrates that there was a difference in accuracy ratings between the immediate and delayed feedback groups at the 'indirect' stage, but they were nearly identical at the direct point. The results given were based on the study of pre-test and post-test data.

The effectiveness of delayed feedback was found to be greater than that of immediate feedback when the feedback was delivered in an indirect manner. Both instant and delayed

feedback are equally helpful when they are delivered through direct feedback.

The study found a notable interaction between the timing of feedback and the type of feedback concerning their impact on the speaking accuracy of Iranian intermediate English as a Foreign Language (EFL) learners. This indicates that the null hypothesis can be rejected.

DISCUSSION

The observed rise in scores among participants in the four groups can perhaps be attributed to a test-retest effect. The participants exhibited enhanced proficiency in the speaking accuracy

tests as a result of their prior exposure to the identical tests or comparable versions of the tests. Nevertheless, asserting that the practice effect was solely responsible for the entirety of the observed rise is an overly broad assertion. The findings of this study align with previous research examining the impact of feedback timing on the improvement of speaking accuracy, as demonstrated by earlier studies conducted by Hunter (2011) and Sheen (2012). These studies provide evidence that second language (L2) development is more effective when feedback is given with a delay. The findings of this study do not suggest that quick feedback is ineffective but rather highlight the higher value placed on delayed indirect input. Doughty (2001) has highlighted the effectiveness of delayed feedback, suggesting that it may be equally as useful as rapid feedback. The findings of Quinn (2014) are also consistent with the findings of the present investigation. The researcher found that delayed feedback exhibited certain advantages in comparison to rapid feedback in his investigation. Siyyari (2005) also discovered a related finding in their study, where they aimed to compare the impact of implicit focus on form through corrective recast with the impact of delayed, explicit focus on form. The researcher observed that, in terms of overall scores, the delayed feedback condition exhibited superior performance compared to the immediate feedback condition, despite the absence of significant differences between the conditions for any specific linguistic component. However, it should be noted that Siyyari's findings may not be entirely comparable to the present study since his instant feedback treatment was implicit, whereas his delayed feedback treatment was explicit.

It is important to note that not all prior studies align with the current study. A study conducted by McDonough and Mackey (2006) examined the reactions of learners towards recasts. The researchers discovered that a phenomenon they referred to as "primed repetition" of recasts, which refers to repetition that occurs not immediately after the recasts but within six turns, was linked to the development of second language proficiency. The act of immediate repetition did

not occur. One possible explanation for this inconsistent result is that the "primed repetition" utilized in their research was more akin to an instantaneous rather than a delayed response. The findings of this study offer more substantiation for certain prior investigations while presenting contrasting outcomes to others. Long (1977) has argued that there are considerations both in favor of and against postponing feedback. Despite the ongoing discourse surrounding the relative merits of immediate and delayed feedback, Mackworth (1950) presents a more compelling argument by asserting that the efficacy of feedback diminishes as the interval between error occurrence and corrective response lengthens. Furthermore, Mackworth posits that when researchers furnish feedback following each erroneous response or failure to respond, participants exhibit heightened vigilance toward their accuracy, surpassing the typical level of attentiveness (p. 201).

This study unveiled a preference for delayed input over quick response. The importance of delayed input should not be disregarded. According to Long (1977), delayed correction offers the benefit of circumventing the inhibitory consequences associated with interrupting learners. Teachers frequently employ role plays or simulations as instructional strategies to facilitate the development of learners' procedural skills in task completion. Interrupting these tasks with quick cognitive feedback may have a detrimental effect on their intended objective. In line with the current investigation, Long (year) elucidates that the term "delayed correction" pertains to the act of correcting after the perceived conclusion of the student's statement (p. 87). The presence of discrepancies between the current study and prior research is not unexpected, as the findings of this study align with the assertions made by Carroll and Swain (1993) and Ellis (2017) regarding the advantageous nature of immediate feedback for learners. Conversely, other scholars, including Lyster and Ranta (1997), Oliver (2000), and Mackey (2003), have posited that delayed corrective feedback is more effective for specific linguistic forms within specific contexts.

CONCLUSION

The findings of the current study regarding the perceived significance of delayed indirect feedback in enhancing speaking accuracy can be viewed as contradictory evidence to the viewpoints of behaviorist theorists such as Brooks (1960), who advocated for minimal time intervals between learners' errors and teachers' provision of correct models. While the current literature on second language acquisition (SLA) has not extensively explored the combined effects of feedback type and timing, existing research on feedback types suggests that delayed feedback, similar to immediate feedback, can support the development of second language (L2) skills.

The findings of this research give empirical support for the effectiveness of delayed feedback, suggesting that it may be equally or even more beneficial than immediate feedback in improving speaking accuracy. The observed inconsistencies may have arisen from variations in the contextual factors of the research, divergent participant characteristics, and disparate forms of feedback. In the study conducted by McDonough and Mackey (2006), it was shown that rapid feedback was equally useful as delayed feedback. However, it is important to note that the delay in feedback observed in their study was not as prolonged as the delay observed in the current study. The findings derived from this research investigation contribute to the existing corpus of literature about the significance of feedback timing and kind in the domain of language acquisition.

The findings of this study indicate that delayed feedback is significantly more efficacious in facilitating indirect feedback compared to direct feedback concerning accuracy. This implies that the consideration of feedback kind should be factored in when determining the appropriate period for providing feedback. Furthermore, according to the findings of the research, it was seen that while delayed feedback proved to be more efficacious for providing indirect feedback, instant feedback also demonstrated effectiveness. Quick feedback was found to be just as beneficial as delayed feedback in the context of providing direct feedback. This implies that both immediate and delayed

feedback can be effectively employed for intermediate learners, contingent upon our intended objectives.

Based on the findings and deductions, feedback can be understood as a complex construct encompassing several abilities, sub-abilities, and elements. Subsequently, it becomes imperative to ascertain if delayed feedback is more favorable than immediate feedback about particular language proficiencies or all aspects of language acquisition. For example, the provision of rapid feedback may be more suitable for lexical repairs, whereas the delay of response may be more appropriate for grammatical errors. Further investigation is required to elucidate this particular aspect in subsequent research endeavors.

References

- Allwright, R. & Bailey, K. M. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge University Press.
- Bartram, M., & Walton, R. (1991). *Correction: Mistake Management*. Language Teaching Publications.
- Brooks, N. (1960). *Language and language learning: Theory and practice*. New York: Harcourt, Brace and Company.
- Chastain, K. (1971). *The development of modern-language skills: theory to practice* (Vol. 14). Philadelphia: Center for Curriculum Development.
- Chaudron, C. (1977). *Teachers' priorities in correcting learners' errors in French immersion classes*. Ontario Institute for Studies in Education.
- Chen, Y. (2019). The effects of immediate versus delayed feedback on writing accuracy. *Language Learning*, 69(3), 407-429.
- Corder, S. P. (1967). The significance of learner's errors. *IRAL-International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. *Georgetown University Round Table on Languages and Linguistics: Educational*

- linguistics, cross-cultural communication, and global interdependence*, 96-108.
- Ellis, R. (2010). A framework for investigating oral and written CF. *Studies in Second Language Acquisition*, 32(2), 335-349.
- Ellis, R. (2017). Oral CF in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (Eds.). *CF in second language teaching and learning: Research, theory, applications, implications* (pp. 3-18). Abingdon: Routledge.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit CF and the acquisition of L2 grammar. *Studies in second language acquisition*, 28(2), 339-368.
- Fanselow, J. F. (1977). The treatment of error in oral work. *Foreign language annals*, 10(5), 583-593.
- Ferris, D., & Hedgcock, J. S. (1998). Teaching ESL composition: purpose. *Process, and Practice*.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be?. *Journal of Second Language Writing*, 10(3), 161-184.
- Fu, T., & Nassaji, H. (2016). CF, learner uptake, and feedback perception in a Chinese as a foreign language classroom. *Studies in Second Language Learning and Teaching*, 6, 161-183.
- Holley, F. M., & King, J. K. (1971). Imitation and correction in foreign language learning. *The Modern Language Journal*, 55(8), 494-498.
- Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal*, 387-398.
- Hughes, R., & Reed, B. S. (2016). *Teaching and researching speaking*. Taylor & Francis.
- Hunter, J. (2012). Small talk': Developing fluency, accuracy, and complexity in speaking. *ELT Journal*, 66 (1), 30-41.
- Karimi, M. N., & Asadnia, F. (2015). EFL teachers' beliefs about oral CF and their feedback-providing practices across learners' proficiency levels. *Journal of Teaching Language Skills*, 7(2), 39-68.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in second language acquisition*, 25(1), 37-63.
- Li, X. (2015). The effects of immediate versus delayed feedback on grammar accuracy. *Applied Linguistics*, 36(3), 351-371.
- Lindsay, C., & Knight, P. (2006). *Learning and teaching English: A course for teachers*. New York: Oxford University Press.
- Long, M. (1977). Teacher feedback on learner error: Mapping cognitions. *TESOL*, 77, 278-294.
- Long, M. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie and T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Lyster, R., & Ranta, L. (1997). CF and learner uptake: Negotiation of form in communicative classrooms. *Studies in second language acquisition*, 19(1), 37-66.
- MacIntyre, P. D., Burns, C., Jessome, A. (2011). Ambivalence about communicating in a second language: A qualitative study of French immersion students' willingness to communicate. *The Modern Language Journal*, 95, 81-96.
- Mackey, A., & Gass, S. M. (2006). Pushing the methodological boundaries in interaction research: An introduction to the special issue. *Studies in Second Language Acquisition*, 28(2), 169-178.
- Park, H. S. (2010). *Teachers' and learners' preferences for error correction*. Unpublished Ph.D Thesis. California State University.
- Praver, M., Rouault, G., & Eidswick, J. (2011). Attitudes and Affect Toward Peer Evaluation in EFL Reading Circles. *The Reading Matrix*, 11(2), 89-101.
- Redmond, M. V., & Vrchota, D. (2006). Everyday public speaking. Allyn & Bacon.
- Robb, T., Ross, S., & Shortreed, I. (1986). The salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20(1), 83-96.

- Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning*, 45(2), 283-331.
- Rolin-Ianziti, J. (2006). Teacher corrective practices in the foreign language classroom: The effect of timing. In C. Bailey & K. Barnett (Eds.), *Proceedings of the Social Change in the 21st Century 2006 Conference*, Queensland University of Technology: Centre for Social Change Research.
- Rolin-Ianziti, J. (2010). The organization of delayed second language correction. *Language Teaching Research*, 14(2), 183-206.
- Rosa, E. M., & Leow, R. P. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics*, 25(2), 269-292.
- Scrivener, J. (2011). *Learning teaching: The essential guide to English language teaching* (3rd ed.). UK: MacMillan.
- Siyyari, M. (2005). *A comparative study of the effect of implicit and delayed, explicit focus on form on Iranian EFL learners' accuracy of oral production*. Unpublished MA thesis, Iran University of Science and Technology
- Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics* 30(2), 510–532.
- Thornbury, S. (2000). Accuracy, fluency and complexity. *Readings in Methodology*, 16, 139- 143.
- Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Varnosfadrani, A. D. (2006). *A comparison of the effects of implicit/ explicit and immediate/delayed CF on learners' performance in tailor-made tests*. Unpublished PhD thesis, University of Auckland.
- Zhang, H., Song, W., Shen, S., & Huang, R. (2014). The effects of blog-mediated peer feedback on learners' motivation, collaboration, and course satisfaction in a second language writing course. *Australasian Journal of Educational Technology*, 30(6)20-33.
- Zhang, Y. (2018). The effects of immediate versus delayed feedback on vocabulary learning. *Journal of Second Language Learning*, 7(1), 1-22.
- Wang, Y. (2017). The effects of immediate versus delayed feedback on speaking fluency. *Language Teaching Research*, 21(3), 299-318.

Biodata

Elham Dehdari is a Ph.D. student of TEFL at Azad University, Sarab, Iran. She has been teaching English for several years at universities and language schools and has published a number of papers in international and national academic journals. Her main areas of interest include teaching methodology and innovative teaching methods.

Email: Elham82dehdari@gmail.com

Nasser Ghafoori is an assistant professor of English Language Teaching at Islamic Azad University, Tabriz Branch, Iran. He mainly teaches language testing, research methodology, and teaching language methodology at the graduate level and his main areas of interest include teacher education, multiple intelligence, language teaching, and research.

Email: mnsaeidi@yahoo.ca

Saeideh Ahangari is an associate professor of English Language Teaching at Islamic Azad University, Tabriz Branch, Iran. She mainly teaches language testing, research methodology, and teaching language methodology at the graduate level and her main areas of interest include teacher education, cooperative learning, language testing, and research. She has published many papers in international and national academic journals and also presented in conferences.

Email: saeideh.ahangari@gmail.com

