Research Paper

The Role of Direct Metalinguistic and Indirect Feedback Timing in Enhancing Willingness to Communicate (WTC) of Iranian EFL Learners

Elham Dehdari[1], Nasser Ghafoori[2]*, Saeideh Ahangari[3]

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

Elham82dehdari@gmail.com

[2]Assistant Professor, Department of English Language, Faculty of Foreign Languages, Islamic AzadUniversity, Tabriz Branch, Tabriz, Iran

ghafoori@iaut.ac.ir

[3]Associate Professor, Department of English Language, Faculty of Foreign Languages, Islamic Azad University, Tabriz Branch, Tabriz, Iran

saeideh.ahangari@gmail.com

ABSTRACT

This study tried to inspect and compare any possible effects of immediate and delayed metalinguistic and indirect feedback on the level of WTC of Iranian EFL learners in task performance. This research was a quasi-experimental work with 120 Participants (63 males and 57 females) who were divided into four groups: 30 students as the direct-immediate feedback group, 30 students as the direct-delayed feedback, 30 students as the indirect-immediate feedback, and 30 students as the indirect-delayed feedback. Different instruments were utilized to collect the required data, including the Oxford Placement Test (OPT), the WTC scale (McCroskey, 1992), and a set of two-way exchange tasks. After homogenizing the participants, based on the OPT test, the participants took the WTC scale as a pretest. As the treatment, the indirect immediate feedback group participants reformulated the errors into the correct form as soon as they committed errors. In the indirect delayed feedback group, the teacher noted the errors and the student's name to provide feedback for that specific student. In the direct metalinguistic immediate feedback group, the participants were provided with both explanations and examples of the correct form of errors they committed. In the metalinguistic delayed group, the teacher wrote notes on the error types and the students' names to provide explanations and correct examples for specific students. The WTC scale was readministered in the post-test and the scores were recorded. Two-way ANOVA was used to analyze the pre-and post-test data. The result of the data analysis showed there was an interaction between the timing and the type of feedback regarding their effect on Iranian intermediate EFL learners' WTC.

Keywords: *Feedback timing, WTC, Meta-linguist feedback, Indirect feedback*

نقش زمان‌بندی مستقیم نراتبانی و بازخورد غیرمستقیم در افزایش تمایل به برقراری ارتباط (WTC) زبان‌آموزان ایرانی زبان انگلیسی در

این مطالعه سعی شده است اثرات احتمالی بازخورد فوری و تأخیری فرازبانی و غیرمستقیم بر سطح WTC زبان آموزان ایرانی در انجام تکالیف بررسی و مقایسه شد. این پژوهش یک کار تجربی با ۱۲۰ شرکت کننده (۶۳ مرد و ۵۷ زن) بود که به چهار گروه ۳۰ دانش آموز به عنوان گروه بازخورد مستقیم فوری، ۳۰ دانش آموز به عنوان بازخورد مستقیم با تأخیر، ۳۰ دانش آموز به عنوان گروه غیرمستقیم تقسیم شدند. بازخورد فوری، و ۳۰ دانش آموز به عنوان بازخورد غیرمستقیم با تأخیر. ابزارهای مختلفی برای جمع‌آوری داده‌های مورد نیاز استفاده شد، از جمله آزمون قرارگیری آکسفورد (OPT)، مقیاس WTC (مک کراسکی، 1992)، و مجموعه‌ای از وظائف مبادله دو طرفه. پس از همگن سازی شرکت کنندگان، بر اساس آزمون OPT، شرکت کنندگان مقیاس WTC را به عنوان پیش آزمون گرفتند. به عنوان درمان، شرکت کنندگان گروه بازخورد فوری غیرمستقیم به محض اینکه مرتکب خطا شدند، خطاها را به شکل صحیح دوباره فرموله کردند. در گروه بازخورد غیرمستقیم، معلم خطاها و نام دانش آموز را یادداشت می کند تا برای آن دانش آموز خاص بازخورد ارائه کند. در گروه بازخورد فوری مستقیم، هم توضیحات و هم نمونه هایی از اشکال صحیح اشتباهاتی که مرتکب شدند به شرکت کنندگان ارائه شد. در گروه تأخیری فرازبانی، معلم یادداشت هایی در مورد انواع خطاها و نام دانش آموزان نوشت تا برای دانش آموزان خاص صحیح را ارائه دهد. مقیاس WTC مجدداً در پس آزمون اجرا شد و نمرات ثبت شد. برای تجزیه و تحلیل داده های پیش آزمون و پس آزمون از آنالیز واریانس دوطرفه استفاده شد. نتایج تجزیه و تحلیل داده‌ها نشان داد که بین زمان‌بندی و نوع بازخورد در رابطه با تأثیر آن‌ها بر WTC زبان‌آموزان ایرانی زبان انگلیسی متوسط تعامل وجود دارد.

کلمات کلیدی : زمان‌بندی بازخورد، WTC، بازخورد فرازبان‌شناس، بازخورد غیرمستقیم

Introduction

It is accepted that learners' meaningful and active involvement is among the most important prerequisites for language learning (Gass, 2003; Long, 1996; Swain, 2005). Communicative approaches in EFL teaching and learning can be regarded as proof of this statement. The communicative teaching approach emphasizes enabling the learners to communicate. It is declared that learners' "willingness to communicate" (WTC) plays an essential role in communication tasks going on in the classroom context. L2 learners who have a high level of WTC seek to communicate more using their L2 in and out of the educational setting (MacIntyre, Baker, Clément, & Donovan, 2003).

Regrettably, due to the lack of knowledge of some teachers in the appropriate use of corrective feedback (CF), some learners lose their WTC and self-confidence. The source of this failure may be the point that feedback in not a unitary concept. It can be provided in different ways including direct and indirect feedback. Previous studies (e.g. Karimi & Asadnia, 2015) have confirmed that there is a difference in the number of employment of CF types, by teachers, depending on how (direct/indirect, explicit/ explicit) and who (teacher, self, peer) provides the feedback. Indeed, feedback-related variables (timing, feedback provider, and type) can lead to different outcomes.

Recently, a bulk of research has focused on the influence of "linguistic", "psychological", and "contextual" factors on learners' L2 WTC (Zarrinabadi 2014; Zarrinabadi & Tanbakooei, 2016). Issues like the effect of instructors' explicit and implicit corrective feedback with regard to their timing on L2 learners' WTC have remained unexplored. In spite of defining the role of WTC in the amount of communication, the WTC degree might be a function of many factors including the way *feedback i*s provided when the learners commit an error. Indeed, a credible answer to the question of CF effectiveness requires that a wider variety of contextual and learner factors in the instructed SLA situation be taken into account (Ellis, 2010).

Based on the way feedback is provided, some types of corrective feedback enhance while some other types lower students' WTC (Macintyre, Burns, & Jessome, 2011). Indeed, the gap in the literature is whether reactive focus on form is more effective when it occurs during the performance of a task or when it is delayed until the task has been completed. It has been discussed in Task Based Language Teaching circles (Ellis, 2017). Moreover, feedback timing is a significant issue since a common position in the academic advice that has been given to teachers is that sometimes feedback should occur immediately and at other times it should be delayed (Hedge 2000; Scrivener, 2011). One further gap in the literature related to the culture-bound nature of previous studies related to feedback and WTC (McCroskey & Richmond, 1987). Indeed, previous studies of WTC and feedback have been generally conducted in contexts other than Iran. Then, there is an urgent need to conduct related studies in the context of Iran. These claims necessitate the need for deeper studies on the role of feedback timing in enhancing EFL learners' WTC across two different feedback types: direct and indirect. Therefore, the present study tries to identify the role of feedback timing in enhancing Iranian EFL learners' WTC across direct and indirect types of feedback.

The results may be useful for language teachers and teacher trainers. English teachers would be justified in the efficiency of immediate and delayed feedback in relation to different kinds of feedback (direct and indirect) provided by teachers. They will find out what type of feedback increases WTC while doing tasks if provided immediately and what type of feedback is more suitable if provided with a delay. Teacher educators can also take advantage of the outcomes of this study. They can raise preservice and in-service teachers' awareness about feedback timing during task performance.

Literature Review

In the instructed SLA research literature, the issue of the timing of feedback has long been marginalized, and interest in it has only recently reemerged (Ellis, 2017). Prior to the communicative language teaching era, the timing of feedback was regularly discussed, with some theorists suggesting that feedback be provided immediately and others that it be delayed. However, since that time, the discussion of feedback timing has greatly diminished.

Beliefs about feedback timing have been significantly influenced by changes in thinking about the way languages are learned. In the 1950s and 60s, theorists urged that error be avoided like sin, and that if it had to be dealt with at all it be dealt with immediately (Hendrickson, 1978). Under the influence of behaviorism, theorists like Brooks (1960) urged that as little time as possible be allowed between the time learners committed errors and the time teachers provided a correct model. Theorists feared the longer learners were exposed to an unaddressed incorrect linguistic behavior, the more likely it was that the behavior would become ingrained. A shift in thinking took hold with Corder's (1967) seminal recognition of the importance of errors as illustrations of the development learners were making with the L2. Teachers were then urged to analyze errors and determine what hypotheses learners might be testing about the L2 through those errors. Fanselow (1977) suggested that this analysis might require novice teachers to delay feedback up to a day to analyze what an error was indicating and how to address it appropriately. Along the same lines, Chastain (1971) suggested that teachers might review common mistakes after communicative activities had ended. Allwright (1975), on the other hand, argued that teachers must not wait but must analyze and provide instant analysis and feedback.

Regarding the existence of contradicting views about the timing of feedback, one factor that might have a role in the timing of feedback is feedback type. Scholars proved that explicit feedback was more influential compared to implicit one (Rosa & Leow, 2004). In the same way, researchers claimed that explicit feedback was more influential than the implicit type (Ellis, Loewen & Erlam, 2006). Nonetheless, some of the other studies (e.g., Fu & Nassaji, 2016) indicated no difference. Just one research (Leeman, 2003) reported that implicit CF was more influential than explicit type. Again, this contradicting finding could be attributed to the timing of feedback. Indeed, no study looks into the role of feedback timing.

Appropriate feedback type and timing can influence the process of language learning through interacting with learners' WTC. The importance of WTC in L2 learning arises from the significant effect that interaction has on acquiring a language, which was proved by linguistic and socio-cultural viewpoints (Kang, 2005). According to Macintyre et al. (1998), WTC which is an important component of modern language pedagogy, can contribute to SLA.

Yashima (2002) discovered that "L2 communication confidence", "perceived communicative competence", and "communication anxiety" are the most important aspects of L2 WTC. Considering Martin and Valdivia's (2017) claim about the influences of feedback type on the level of anxiety among EFL learners, it can be hypothesized that there is an interaction between feedback type, feedback timing, and WTC among Iranian EFL learners. Therefore, considering the existing gap in the literature on the one hand and the significance of conducting it on the other hand, this study was going to find answers to the following research questions:

RQ1: Does feedback timing have any significant effect on intermediate EFL learners'   WTC in task performance?"

RQ2: Does feedback type have any significant effect on intermediate EFL learners' WTC in task performance?"

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

Method

Participants

This study began in the summer of 2019 in Iranmehr language institutes in Tehran, Iran. To accomplish this research, the researcher recruited students at the intermediate level as participants. The classes were held for 16 sessions, twice a week. Using a quasi-experimental research design, the researcher gathered 120 EFL learners from a population of 200 students who were selected based on their 'Oxford Placement Test (OPT) scores. So, 30 EFL learners were selected as the first experimental group (direct-immediate feedback), 30 students were selected as the second experimental group (direct-delayed feedback), 30 participants were selected as the third experimental group (indirect-immediate feedback), and 30 participants were selected as the fourth experimental group (indirect-delayed feedback). The students were selected from male (n= 63) and female (n= 57) genders. The participants' age range was from 15 to 31 (m= 21.5). The same teacher was asked to teach all students. The logic for choosing the same teacher for all participants was to neutralize any variable that might result from teachers' personal factors. Three students were dropped during the process of intervention and data gathering because their WTC questionnaire was not filled completely.

The participation in the study was not voluntary in order not to impair the external validity through contaminated sampling; because, as Dornyei (2007, p.100) refers to what he calls the "problem of self-selection" volunteer participants may be a more motivated or "eager beaver" in research and this may damage the generalizability of the study. It needs to be added that for this study intact classes were used to lower the role of intervening variables as much as possible.

Materials and Instruments
Oxford Placement Test (OPT)

To select homogeneous EFL learners, the OPT test (version I) was used. "Oxford University Press and the University of Cambridge Local Examinations Syndicate" have prepared this English language test. The results are used to show the ability of participants to use English for communication purposes. The test includes 60 items about language structures in multiple-choice formats. As test designers declared, participants, whose score is between 37 and 46, are at the intermediate level. The allowed time to answer is 30 minutes.

Speaking Accuracy Test

With the purpose of checking immediate and delayed feedback effects on participants' speaking accuracy, the researcher adopted a speaking accuracy test from the PET test (2018). To this end, the participants were required to perform a task with a partner in 4 minutes. The voices were recorded for future analysis. Since the target structure was the use of English tense by students, the number of accurate verbs was divided by the total verbs in their speech. This ratio was regarded as the participants' accuracy test score (Yuan & Ellis, 2003). The test had been checked for its reliability, validity, and practicality in the context of Iran by previous researchers (e.g. Assasi, 2018). To ascertain the reliability, two raters rated the speaking accuracy tests. Therefore, reliability was checked through interrater reliability. The validity of the measure was also confirmed through panel discussion. Indeed, a panel of TEFL experts was asked to give their view of the test about its validity.

WTC Scale

McCroskey (1992) proposed The WTC scale for gathering data about participants' WTC. This instrument included 20 items. These items are about four communication settings (including "public speaking", "talking in meetings", "talking in small groups", and "talking in dyads") and

three kinds of receivers: acquaintances, friends, and strangers. The participants were supposed to write the percentage of time (each set from 0 to 100) they would select to communicate. According to McCroskey and Richmond (1990), the internal reliability of the WTC scale has been estimated to be above 0.90. To use the questionnaire, the scale was piloted on 20 students and the results were analyzed for reliability; the results were satisfactory.

Language Tasks
During the treatment, participants performed one task in each session. A two-way exchange task (Spot-the-differences) that elicits collaborative interaction was selected for this study. Table 1 presents some examples of the given topics.

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

Topic	Task type	Description of the task
Persian and 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)

Spot the Differences

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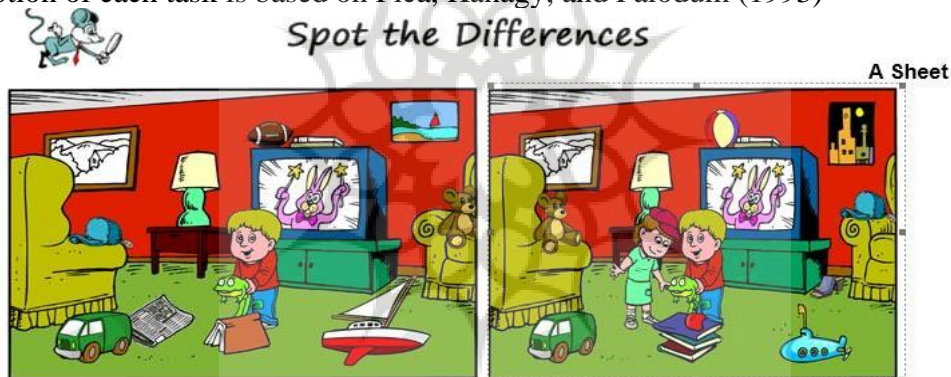
Describe your picture to your partner and try to find as many differences as you can. When you are finished compare your pictures and write down the differences that you found.

"These two-way exchanges promote negotiation of *meaning* and *form*" (Skehan, 2003, p 400). An example of goal-oriented tasks for EFL learners is to perform in pairs and find 15 differences in two pictures of a room. To accomplish this task, the participants may require specific vocabularies and grammatical points to reach to a shared understanding.

Procedure
After selecting the final participants of the study based on the OPT test, they took the speaking accuracy test by performing a task as the pretest. To this end, the participants were required to do a task with a partner in 4 minutes. The voices were recorded for future analysis. As the target structure of the current research was to use the English tense by students, the number of accurate verbs was divided by the number of total verbs in their speech (Yuan & Ellis, 2003). The responses were recorded to be analyzed for accuracy. Moreover, the WTC scale was used and the scores were recorded. Then, they were assigned to each of the four groups based on the type of feedback they were going to receive (elicitation and elaboration) and the timing of feedback (immediate and delayed).

In the elicitation immediate feedback group, the participants were required to reformulate the errors into the correct form as soon as they committed errors. As maintained before, the errors in this group were corrected during the task performance. An instance of the error correction process in this group is provided below:

Example 1:
S1: Just this? Can you tell me another difference?
S2: Mm... yeah there is, the table is* not in the right picture
T: Tables IS?
S: ARE I think! The tables are not in the right picture.
(It needs to be added that the correction was without delay, just like a conversation turn)

In the elicitation-delayed feedback group, the teacher noted down the errors and the participants' names to provide the feedback for that specific student. For the feedback to be considered delayed, it was provided after the task was finished (Li, Zhu, & Ellis, 2016). Indeed, after the peers' interaction was finished, the teacher used to start discussing the errors. As maintained before, the errors in this group were corrected after the task performance. An instance of the error correction process in this group is provided below:

Example 2:
 [The task is finished; a student uttered the incorrect sentence]:
S: The boy and the girl have cut their hair
T: You [pointing to the student who committed the error] said: The boy and the girl have cut their hair? The GIRL (raising one of his fingers) and the
BOY (raising the second finger), are two. Then, we say: The boy and the girl...?!
S: Have?
T: Excellent, so, the boy and the [...]
S: [...] girl HAVE cut the hair

In the elaboration immediate feedback group, the participants were provided with both explanations and examples of the correct form of errors they committed (Nassaji, 2015). As maintained before, the errors in this group were corrected during the task performance. An instance of the error correction process in this group is provided below:

Example 3: [the student is describing a picture]
S1: what about the wall? Don't you notice anything?
S2: The Window... the glass of the window is broken*
T: No! The break is an irregular verb; we do not add "ed" to make irregular verbs passive. We need to memorize them as they are. There are other irregular verbs like...
THE CLASS: go (went), come (come), get (got), etc.
T: Yes, then we say: the window is BROKEN; break ⟶ broken

In the last group, the elaboration delayed group, the teacher used to write notes about the type of errors and the names of students in order to provide explanations and correct examples for specific students. As mentioned above, for the feedback to be considered delayed, it was provided after the task was finished (Li, Zhu, & Ellis, 2016). Indeed, after the peers' interaction was finished, the teacher used to start explaining the errors. An instance of the error correction process in this group is provided below:

Example 4:
[The task is finished; one of the students' errors was the sentence: Does she have* a flower in her hand?]
T: [pointing to the student who committed the error] When you want to make questions using "does", we must use "have" instead of "has"; for example, DOES Ali HAVE a car? You said: Does she have* a flower in her hand? then, you must say: Does she...!

S: HAVE a flower in your hand?

After the treatment was finished, the participants took another speaking accuracy test as a posttest and refilled the WTC scale. It needs to be added that the focus of the accuracy component of speaking task performance was on the use of English tenses.

Data analysis

The answers to the speaking accuracy test, both for pre- and post-test, were measured, scored were fed into SPSS 24 software. The answers to the WTC questionnaire were also entered into SPSS. Reversed scores were computed and data were independently checked for entry errors. Descriptive analysis was conducted to report the mean, standard deviation and number for the pretest and post-test scores. Besides, the researcher ran the Smirnov-Kolmogrove test to examine the homogeneity of the scores in all four groups. As the results were acceptable, the researcher compared the means of the test scores inferentially at $p = .01$ level. Then, a Two-way ANOVA test was run for pre- and post-tests to check the differences among the groups.

Results

To measure the effects and the interaction of the two independent variables on the WTC level, there was a need to compare pre-and post-test WTC 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 WTC difference among the four groups, before the treatment. Table 2 presents the descriptive statistics (e.g., mean and SD) for each combination of the independent variable group. Furthermore, the following table presents "Total" rows, which allows 'means' and 'standard deviations' for groups only split by only one independent variable.

As Table 2 illustrates, at the beginning of the research the four groups were roughly the same in terms of their WTC scores. According to Table 2, the highest WTC scores belonged to the delayed indirect group members (M= 0.541) while the lowest scores belonged to the delayed direct group participants (M= 0.535).

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

Descriptive Statistics

Dependent Variable: WTC pretest

Feedback type	Feedback timing	Mean	Std. Deviation	N
indirect	immediate	.5393	.07446	30
	delayed	.5417	.09359	30
	Total	.5405	.08386	60
direct	immediate	.5413	.08216	30
	delayed	.5357	.07133	30
	Total	.5385	.07633	60
Total	immediate	.5403	.07774	60
	delayed	.5387	.08255	60
	Total	.5395	.07985	120

Although the mean scores of WTC were different in the four groups, the statistical significance of the differences needed to be checked. Indeed, the score differences could have

been insignificant. Since there were two factors with two levels, Two-way ANOVA was used to compare the results. After it was assured that the six assumptions of the Two-way ANOVA test had been met, the four groups' pretest scores in WTC were compared. The results related to the Two-way ANOVA test are presented below in Table 3.

The particular rows to be considered are the "Feedback type pre", "Feedback timing pre" and "Feedback type pre * Feedback timing pre" rows. These rows reveal if the independent variables (the "feedback type" and "feedback timing" rows) and their interaction (the "feedback type*feedback timing" row) do not have any significant difference in mean WTC scores.

Table 3
Two-way ANOVA results

Tests of Between-Subjects Effects

Dependent Variable : WTC pre-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.001[a]	3	.000	.035	.991
Intercept	34.927	1	34.927	5.344E3	.000
Feedback type	.000	1	.000	.018	.892
Feedback timing	8.333E-5	1	8.333E-5	.013	.910
Feedback type * Feedback timing	.000	1	.000	.073	.787
Error	.758	116	.007		
Total	35.686	120			
Corrected Total	.759	119			

a. R Squared = .001 (Adjusted R Squared = -025)

The "Sig." column shows that there is not any statistically significant interaction at the $p = .78 > 0.01$, meaning that, there is no significant difference among all groups. Simple main effects analysis also showed that neither feedback timing nor feedback type were significantly different (feedback type: $p = .89 > 0.01$; feedback timing: $p = .91 > 0.01$). To sum it up, it is induced that there was no significant difference in mean WTC score of all groups. For better understanding, the similarity of WTC scores in the pretest, the mean "WTC" score plot for each combination of groups of "feedback type" and "feedback timing" are showed below as

Figure 1
Plot of pretest WTC results
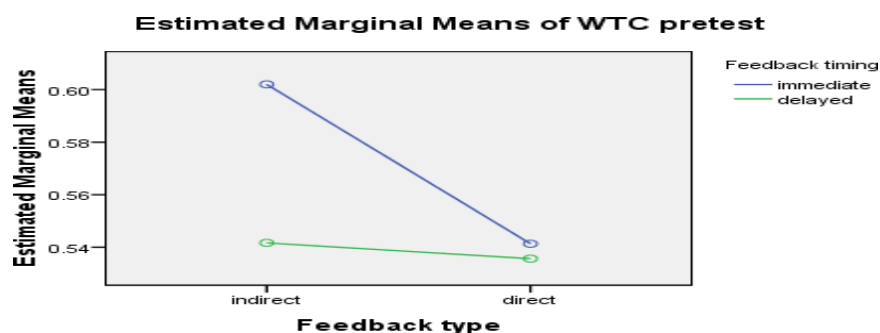
As generally observed in the figures, usually, an interactional effect can be shown in a group of non-parallel or crossing lines (both lines are started from top-left to down-right). However, it can be seen from this graph that the mean difference of WTC at 'indirect' and 'direct' are not substantial.

Up to this point, it has been showed that there was not any significant difference among the four groups' WTC levels before the treatment starts. In order to measure the effects of treatment on the WTC, there was a need to statistically test the significance of the difference of posttest WTC scores across the four groups. Again, since there was one dependent variable and two independent variables, the statistical difference should have been checked through Two-way ANOVA. Similar to what was done for the pretest scores, the mean WTC scores needed to be measured to see if they were different across the four groups or not. Indeed, before running Two-way ANOVA test for checking the significance of the difference, it should be checked if there are any differences or not.   To this end, descriptive statistics, as showed in Table 4, were run to notice if there was any WTC difference among the four groups after the treatment or not. Table 4 offers the mean and standard deviation for each combination of the groups of the independent variables. Moreover, the following table offers "Total" rows, allowing means and standard deviations for groups only split by one independent variable, or none, to be known.

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

Feedback type	Feedback timing	Mean	Std. Deviation	N
indirect	immediate	.5350	.09339	30
	delayed	.5807	.05878	30
	Total	.5578	.08072	60
Direct	immediate	.5420	.06178	30
	delayed	.5800	.07488	30
	Total	.5610	.07070	60
Total	immediate	.5385	.07859	60
	delayed	.5803	.06674	60
	Total	.5594	.07557	120

As Table 4 illustrates, after the treatment, the four groups were not the same in terms of their WTC scores. According to Table 4, the highest WTC scores belonged to the indirect delayed group members (M= 0.5807) while the lowest scores belonged to the indirect immediate group participants (M= 0.535). Although the mean scores of WTC were different in the four groups, the differences needed to be tested for statistical significance. Indeed, the score differences could have been insignificant. Therefore, to explore the statistical significance of the score differences, Two-way ANOVA needed to be run. As maintained before, the six pre-assumptions related to Two-way ANOVA have been met. Again, the first four assumptions are met by the design of the study; however, the last two assumptions were checked statistically.

After the pre-assumptions of the Two-way ANOVA were considered and the mean scores of posttest WTC across the four groups were different, the Two-way ANOVA test was run and the findings related to it are shown in Table 5.

Table 5

Two-way ANOVA analysis results

Dependent Variable: WTC post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Square
Corrected Model	.053[a]	3	.018	3.286	.023	.078
Intercept	37.554	1	37.554	6.954E3	.000	.984
Feedback type post	.000	1	.000	.056	.814	.000
Feedback timing post	.053	1	.053	9.722	.002	.077
Feedback type post * feedback timing post	.000	1	.000	.082	.776	.001
Error	.626	116	.005			
Total	38.233	120				
Corrected Total	.680	119				

a. R Squared = .078 (Adjusted R Squared = .055)

The required rows to be reported from this table were the "Feedback type post", "Feedback timing post" and "Feedback type post*Feedback timing post" rows. These rows show if the independent variables ("feedback type" and "feedback timing" rows) and their interaction (the "feedback type*feedback timing" row) have a significant difference of mean WTC scores.

According to the "Sig." column, there was a significant difference at the $p = .01$ level regarding the feedback timing. Therefore, there is a difference between immediate and delayed groups regarding their posttest WTC score (p= 0.002<0.01). Moreover, it can be noted that there was no significant effect regarding the feedback type (p= 0.814>0.01) and the interaction between feedback timing and type (p= 0.776>0.01). This means that there was not any difference between direct and indirect feedback groups with regard to their posttest WTC levels. To sum it up, it can be inferred from the above table that there was a difference in the mean WTC score across all groups.

Table 5 presents the effect size as well as the significance of the difference. To find out the amount of mean WTC score variance that has resulted from the independent variable, effect size, i.e., the last column of the above table (partial Etta squared) needs to be reported. Regarding feedback timing, an effect size of 0.07 has been reported. This means that 0.07 of the variances across the four groups has been due to the timing of feedback. However, it needs to be added that eta square estimates only the effect size in the sample not in the population.

For better understanding, the difference of WTC scores in the posttest and the interaction between feedback type and feedback timing, the plot of the mean "WTC" scores for each group combination of "feedback type" and "feedback timing" are shown below as Figure 2.

Figure 2

Estimated Marginal Means of WTC post test
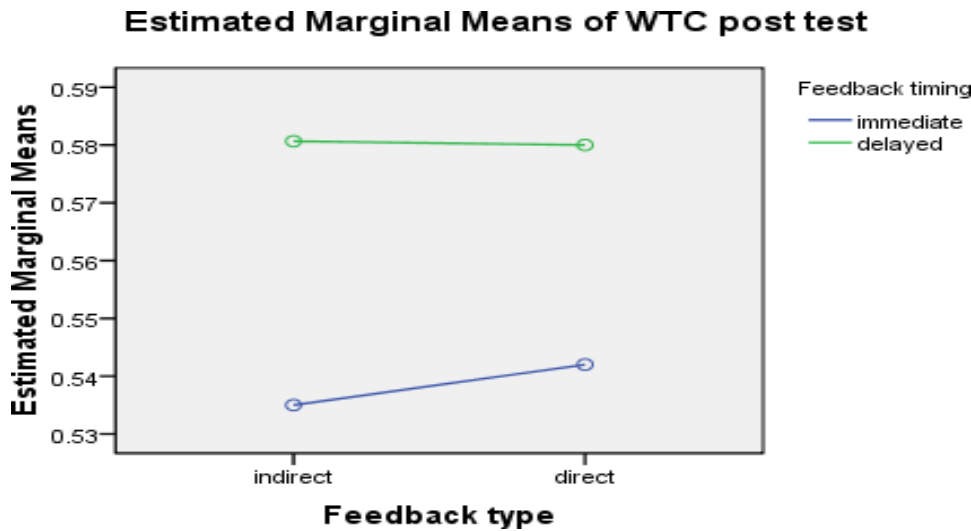
While the estimated marginal means (EMM) of immediate and delayed feedback are significantly different at the indirect point (immediate EMM: 0.6, delayed EMM: 0.7), they are similar at the direct point (around 0.6). An interaction effect can usually be seen as a set of non-parallel lines. As can be seen in this plot, the lines are not actually crossing at a 'direct' point. Simply put, the graph shows that while WTC scores of immediate and delayed feedback groups are different at the 'indirect' point, they are almost equal at the direct point. Then, based on the pre-and post-test data analysis, the following results were obtained:

- Delayed type of feedback was more effective compared to immediate feedback while the data was provided through indirect feedback.
- Immediate feedback and delayed feedback were both effective when the data was provided through direct feedback
- There was a positive interaction between feedback timing and feedback type considering their influence on EFL learners' WTC level, so, it can be said that the hypothesis is rejected.

Discussion

The analysis of the findings indicated that immediate and delayed feedback had a significant difference regarding their effect on participants' WTC. Simply put, while the mean WTC scores of the two delayed feedback groups were 0.58, those of the immediate feedback groups were 53.5 and 54.2. These results highlight the significance of feedback timing in relation to WTC. The effectiveness of delayed feedback over immediate feedback can be justified by the concept of anxiety involved in immediate feedback. MacIntyre (2007) maintains that language anxiety and motivation are defining factors in willingness to communicate in L2. Considering Martin and Valdivia's (2017) claim that language anxiety is correlated with the type of feedback given to the learners and MacIntyre's (2007) claim about the role of anxiety in WTC, the priority of delayed feedback over immediate feedback can be described by the point that language learners may have lower levels of anxiety when the feedback is provided in delayed form.

A study with contradicting results was done by Ghahari and Piruznejad (2016) which set out to examine the effects of recast (indirect) and direct feedback on young EFL learners' WTC. The obtained results revealed that the recast group showed more WTC. They concluded that for young learners the use of a less direct way might be more effective for WTC. However, the

obtained results of the current research show no significant difference between direct and indirect feedback concerning enhancing EFL learners' WTC. This finding can be justified by the age of the learners. Indeed, feedback might not influence different age groups in a similar way. This claim is made because McCroskey, Richmond, Daly, and Falcione (1977) reported that there is a relationship between age groups and communication apprehension and self-esteem. Considering McCroskey's (1992) claim that communication apprehension is correlated with WTC, it can be perceived that the contradiction between Ghahari and Piruznejad (2016) and the present study findings is due to the age of the participants.

Although the results of this study suggested the priority of delayed feedback over immediate feedback, not all other previous research obtained the same results. For instance, Zadkhast and Farahian (2017) who were set out to measure the immediate and delayed feedback effect on EFL learners' WTC level. Although the findings suggested the positive effect of immediate and delayed corrective feedback on EFL learners' WTC, immediate feedback was revealed to be more effective. This contradiction can be justified by the two existing differences in the nature of participants. First, only females were the participants of Zadkhast and Farahian's (2017) study. Second, the participants in the study conducted by Zadkhast and Farahian (2017) were not homogenized. Considering Gholami (2015) found a difference between male and female learners regarding their cognitive as well as social traits, the contradicting findings of this study with those of Zadkhast and Farahian (2017) can be justified by the gender of the participants. Indeed, Gholami (2015) maintained that there is a correlation between gender and WTC.

The observed inconsistencies in this study could have been due to differences in the context of the studies, different characteristics of the participants, and different types of feedback. For instance, McDonough and Mackey's (2006) study found that both immediate and delayed feedback is effective, but in this study, the delayed feedback was not as influential as immediate. The obtained results would add to the body of literature related to the role of feedback timing and type in the social and psychological state of the learners.

Considering the obtained findings, one point needs to be added here: there is an interaction between feedback timing and type among intermediate Iranian EFL learners who are adults. In the same way, compared to immediate feedback, delayed feedback was observed to be preferred and more effective. Moreover, the effectiveness of the delayed feedback was more powerful when the feedback was provided in the form of indirect feedback. The focus on participants, context (EFL), proficiency level, and age of the participants lies in the fact that these factors might interact with the degree to which feedback timing and type influence speaking ability. For example, elementary participants might be different from pre-intermediate and intermediate participants in terms of the way immediate and delayed feedback may influence them. MacIntyre et al (2001), who propose a theoretical empirical model to illustrate the individual and contextual variables' role in WTC by combining motivational and attitudinal factors, support this point. They considered WTC as an effective situational variable in the four language skills. Peng (2007) mentioned that "various social-psychological, linguistic, and communicative variables as precursors of L2 communication" (p. 545).

Conclusions

The findings indicated that the use of delayed or immediate feedback is not a yes-or-no matter. The results also revealed the effectiveness of delayed feedback for indirect feedback compared to direct feedback regarding the WTC. This suggests that the type of feedback should be taken into account while deciding on feedback timing.

Considering the existence of some limitations and delimitations in the process of conducting this study provided the researcher with fresh ideas that could be insightful for the researcher in this field.

- First, the findings showed the priority of delayed feedback over immediate feedback in developing intermediate students' WTC, when the correct form is indirectly provided for the learners. The usefulness of delayed feedback highlights the hypothesis that feedback type has a defining role. Then, future researchers are suggested to repeat this research on a wider range of feedback types.

- Second, the obtained results signified that delayed feedback was more effective than immediate feedback. Other researchers to give us a better understanding of the reasons can conduct a qualitative study. Simply put, future researchers are suggested to dig into the reasons for the priority of delayed feedback over immediate feedback.

- Third, this study recruited participants aged between 15 and 30. Moreover, the mean age of the immediate groups was around 21.4 while that of the delayed groups was around 23.5. Since many social, affective, and psychological personal traits might change as an individual grows up, future researchers are suggested to conduct the present study across different age groups to see if in other ages similar findings would be reported or not.

- The present study has focused on the interaction between two feedback types and two feedback timings regarding WTC performance. Needless to say, language is a multifaceted concept, consisting of different skills, subskills, and components. Then, there is a need to check if delayed feedback is preferable to immediate feedback regarding specific language abilities or all language aspects. For instance, lexical corrections may lend themselves to immediate feedback while grammatical corrections might better be delayed. This point needs to be clarified in future studies.

- An examination of the influences of feedback timing on speaking or writing fluency and can be insightful. Fluency is especially focused since higher levels of WTC might raise fluency in a significant way (Yousefi & Kasaian, 2014).

- Finally, future researchers can study the effects of feedback type and timing across different proficiency levels.

- The obtained findings can be useful for stakeholders in the field of language teaching and learning. Then, future researchers are suggested to conduct the present study on a wider range of feedback types.

- The findings have some implications for language teachers, language learners, materials writers, syllabus designers, and test developers.

- First, the findings suggest that the type of feedback can have a significant influence on Iranian EFL students' speaking accuracy or WTC and teachers can feel confident about providing either immediately or after a delay. However, delayed feedback might be more effective, especially when eliciting the correct form from EFL learners. This enhances both WTC and the accuracy of speaking. This is the same as many results from previous studies that have compared and contrasted the influences of immediate feedback and delayed feedback.

- The use of delayed or immediate feedback is not a yes-or-no matter. The findings of this research revealed that delayed feedback is much more effective for indirect feedback than direct feedback regarding WTC. This suggests that the type of feedback should be considered while deciding on feedback timing.

- According to the results of this research, although delayed feedback was more effective for indirect feedback, immediate feedback was significant, too. Even immediate feedback was as effective as delayed feedback when direct is concerned. This suggests that both types of feedback can be utilized effectively for intermediate learners depending on our purpose.

References

Allwright, R. (1975). Problems in the study of the language teacher's treatment of learner error. In M. Burt & H. Dulay (Eds.), On TESOL '75: *New directions in language learning, teaching, and bilingual education* (pp. 96–109). Washington, DC: TESOL

Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language learning, 50*(2), 311-341.

Cao, Y. (2011). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System, 39*(4), 468-479.

Cao, Y. (2014). A sociocognitive perspective on second language classroom willingness to communicate. *TESOL Quarterly, 48*(4), 789-814.

Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System, 34*(4), 480-493.

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.

Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of language and social psychology, 22(*2), 190-209.

Corder, S. P. (1967). The significance of learner's errors. *IRAL-International Review of Applied Linguistics in Language Teaching*, *5*(1-4), 161-170.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA review, 19*(1), 42-68.

Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. *Georgetown University Round Table on Languages and Linguistics: Educational linguistics, crosscultural communication, and global interdependence*, 96-108.

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

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition, 32*(2), 335-349.

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (Eds.). *Corrective feedback 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 corrective feedback 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.

Freiermuth, M., & Jarrell, D. (2006). Willingness to communicate: can online chat help? *International journal of applied linguistics, 16*(2), 189-212.

Fu, T., & Nassaji, H. (2016). Corrective feedback, learner uptake, and feedback perception in a Chinese as a foreign language classroom. *Studies in Second Language Learning and Teaching, 6,* 161–183.

Gass, S. (2003). Differential effects of attention. *Language learning, 53*(3), 497-546.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.

Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern language journal,* 387-398.

Holley, F. M., & King, J. K. (1971). Imitation and correction in foreign language learning. *The modern language journal*, *55*(8), 494-498.

Kang, S. J. (2005). Dynamic Emergence of Situational Willingness to Communicate in a Second Language. *System, 33*, 277-292.

Karimi, M. N., & Asadnia, F. (2015). EFL teachers' beliefs about oral corrective feedback 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, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal, 100*(1), 276-295.

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). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in second language acquisition*, *19*(1), 37-66.

MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The modern language journal, 91*(4), 564-576.

MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of language and social psychology, 15*(1), 3-26.

MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in second language acquisition, 23*(3), 369-388.

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.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562.

MacIntyre, P., Baker, S., Clément, R., & Donovan, L. (2003). Talking in order to learn: Willingness to communicate and intensive language programs. *Canadian Modern Language Review*, *59*(4), 589-608.

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.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Researches on the Measurement of Human Performance, 21*(2) 56-79.

McCroskey, J. C. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly, 40*(1), 16-25.

McCroskey, J. C., & Richmond, V. P. (1987). Willingness to communicate and interpersonal communication. *Personality and interpersonal communication, 2, 129-*156.

McCroskey, J. C., & Richmond, V. P. (1992). Increasing teacher influence through immediacy. Power in the classroom. *Communication, control, and concern, 45*, 200-211.

McCroskey, J. C., Richmond, V. P., Daly, J. A., & Falcione, R. L. (1977). Studies of the relationship between communication apprehension and self-esteem. *Human communication research, 3*(3), 269-277.

Nassaji, H. (2015). The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice. London: Bloomsbury.

Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language learning, 60*(4), 834-876.

Peng, K. R. (2010). An exploratory cross-sectional study of inter language pragmatic development. *Studies in Second Language Acquisition, 22,* 27-67.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. *Multilingual Matters, 9*(2) 45-66.

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.

Skehan, P. (1989*). Individual differences in second language learning*. London: Edward Arnold.

Spada, N. (2006). Communicative language teaching: Current status and future prospects. In J. Cummins & C. Davison (Eds.), *Kluwer handbook of English language teaching* (pp.259 276). Amsterdam: Klewer Publications

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 495-508).

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86*(1), 54-66.

Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and effect on willingness to communicate and second language communication. *Language learning, 54*(1), 119-152.

Zadkhast, M., & Farahian, M. (2017). The impact of immediate and delayed corrective feedback on Iranian EFL learners' willingness to communicate. *International Journal of Applied Linguistics and English Literature, 6*(6), 28-39.

Zarrinabadi, N. (2014). Communicating in a second language: Investigating the effect of the teacher on learners' willingness to communicate. *System, 42,* 288-295.

Zarrinabadi, N., & Tanbakooei, N. (2016). Willingness to communicate: Rise, development, and some future directions. *Language and Linguistics Compass*, *10*(1), 30-45.

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.