The Effect of Metalinguistic Corrective Feedback on EFL Learners' Grammatical Accuracy

Somaieh Abdollahzadeh
Azad University of Miandoab, Iran

Abstract—In everyday classrooms, teachers and their students must interact with each other on many levels. High communication abilities are required in order to teach and motivate these students effectively. Therefore, feedback is an important part of the teaching and learning. There are different kinds of feedback. The present study was designed to investigate the effect of metalinguistic feedback on grammatical accuracy among Iranian L2 learners. Sixty learners from Sahand language Institute in Miandoab after taking grammatical judgment test which was administered to homogenize them, were placed in two control and experimental groups. The experimental group, received metalinguistic feedback as the treatment. The researcher corrected the learners errors related to past tense and pronoun after they finished the retelling according to principles of Jigsaw task. But learners in control group did not receive any feedback. After treatment, which lasted for six sessions, post-test was given to both control and experimental groups to observe the difference resulted from the treatment. To be sure about the significance of the difference between post-test means of both groups, a t-test was used. The results at the end showed that learners in experimental group outperformed control group. After that, other tests (pronoun and tense tests) were given to the learners in both control and experimental groups. The data collected was computed through t-test which revealed that the effect of metalinguistic feedback on accuracy of tense is greater than pronoun. The findings of this study can be helpful for language teachers and teacher trainers.

Index Terms—feedback, metalinguistic feedback, task, accuracy

I. INTRODUCTION

A. Background

The term "corrective feedback" has recently been imperative part of the foreign language teaching. As its name refers, corrective feedback is used to give information on correctness of learner's utterances and provide them with the correct form of their erroneous production.

Kepner (1991 as cited in Grami, 2005) defined feedback in general as "any procedures used to inform a learner whether an instructional response is right or wrong." (p. 141)

According to Lyster and Ranta (1997) there are different kinds of feedback: explicit correction, recast, clarification, metalinguistic feedback, elicitation, and repetition. It is the teachers' task to decide which kind of feedback to be used. However, this decision should be based on the teachers' regards as to which type of feedback can bring the most effective outcome.

The present study tries to describe the effect of metalinguistic feedback on L2 learners' grammatical accuracy. The researcher wants to know if using metalinguistic feedback would lead to a better chance of grammatical accuracy occurring.

B. Research Questions

This research seeks to answer the following questions:

Research question 1: Does metalinguistic corrective feedback have any effects on the grammatical accuracy of EFL learners?

Research question 2: Does metalinguistic corrective feedback have any effects on the accuracy of tense?

Research question 3: Does metalinguistic corrective feedback have any effects on the accuracy of pronoun?

II. REVIEW OF THE RELATED LITERATURE

A. Theoretical Background

According to Brandet (2008) information given to learners related to some aspects of their performance on a task is called feedback. He maintains that teachers provide this piece of information for learners and there should be a balance in the authenticity of feedback provided for learners and also teacher feedback must be less face threatening.

According to Naeini (2008), an effective instruction entails feedback. As he argues, in classes where learners are deprived of corrective feedback, no pedagogically significant result is obtained. He believes that learning happens when there is an exchange of information between learners and teachers. Furthermore, he claims that there should even be a considerable amount of interaction within learners.

Feedback is an essential part of the teaching and learning process. There are many kinds of feedback and many kinds of people. Since students are different in personality and trait, the type of feedback given will either positively or negatively influence that student. Thus, it is important to discuss the different kinds of feedback and their influences on students. Feedback is an imperative part for learners in language learning.

The purpose of feedback is to teach skills that help learners to improve their writing skill. (Williams, 2003).

The most important thing while giving feedback is adopting a positive attitude toward learner errors. If the learner receives only negative feedback, he/she may easily be disappointed from trying to learn complex structures. However, feedback sessions can be an important experience for the student if the teacher shows the strong points as well (Gulcat & Ozagac, 2004)

Kepner (1991 as cited in Grami, 2005) defined feedback in general as "any procedures used to inform a learner whether an instructional response is right or wrong." (p. 141)

It is important to consider that, in most studies, the effectiveness of corrective feedback are only assessed in terms of learners' immediate response to the feedback (Ellis,2001).The learners' responses to feedback can not be equated with ultimate use of feedback in real life situation, that is, in artificial classroom context, learners may notice teacher's feedback and sometimes could produce the correct form of the first utterance, but it does not mean that learner will never commit that kind of error again. Therefore, more studies are needed to provide supportive evidence for effectiveness of feedback (Ahmad Shah, 2003).

B. Different Kinds of Feedback

Lyster & Ranta (1997) distinguished six types of feedback, namely **explicit correction, recasts, clarification requests, metalinguistic feedback, elicitation, and repetition**.

Explicit correction: clearly indicating that student's sentence was wrong and the teacher provides the correct form.

Recast: Recast is a kind of corrective feedback in which teachers without directly showing that the student's utterance was incorrect, they implicitly reformulates the student's error, or provides the correction.

Clarification: Teachers by using phrases like "excuse me" or "I don't understand" show that the message has not been understood or that the student's utterance includes some kind of error and that a repetition or a reformulation is required.

Metalinguistic Clues: In this kind of feedback the teacher poses questions or provides comments or information related to the learner's utterance.

Repetition: the teacher repeats the student's error and adjusts intonation to draw student's attention to it.

C. Task Based Language Teaching

Task based language teaching is an approach to the design of language course in which the end is not a list of linguistic items, but a collection of task (Nunan, 1999). According to Harmer (2001), task- based language teaching refers to an approach based on the use of tasks as the core unit of planning and instruction in language teaching. The idea of task-based lesson was mostly popularized by prabhu who speculated that students were just likely to learn language if they were thinking about a non-linguistic problem than if they were concentrating on particular language forms.

Task-based language teaching is generally characterized as a development within the communicative approach. It has resulted in response to some disadvantages of traditional PPP approach represented by the procedure of presentation, practice and performance focusing language learning as a developmental process which develops communication and social interaction (Ellis, 2003).

Ellis (2003) believes that task-based language teaching constitutes a strong version of CLT, and tasks provide the basis for an entire language curriculum.

Johnson and Johnson (1998) manifested that in TBLT much interest is focused on the nature of classroom activities (tasks) learners are asked to undertake, and on the possibility of using these tasks as the basis for syllabus design. They refer to Prabhu's work in 1987 on the procedural syllabus as an important attempt at task-based teaching.

Tasks are everyday activities which need language, for example writing a letter. As well as activities that can be done without resorting to any language, e.g; painting a door (Long, 1985 as cited in Ellis, 2003). Task is an activity where the focus is put on meaning, communicative problems are solved, real world activities are related to each other and the assessment of the task is on the basis of outcome (Nunan, 1989).

1. Task types

According to different scholars there are many different kinds of task. Pattison (1987, as cited in Nunan, 2003) sets out seven task types:

● **Questions and answers**: Question and answer tasks based on the notion of creating an information gap which lets the learners make a personal choice from a list of language items and they are to discover their classmate's choice.

● **Dialogs and role play:** In theses kinds of activities if learners are given some choice of what to say, and if there is a clear aim to be achieved, learners will participate willingly and learn much better than that when they are told to repeat the dialog, because more repetition sometimes seems to be a tiring action**.**

● **Matching activities:** In these kinds of tasks the learners' task is to match the items or to complete pairs.

● **Communication strategies:** These are activities which help learners practice communication strategies like asking for feedback.

● **Picture stories:** They create communication activities through the use of pictures.

● **Puzzles and problems:** These kinds of tasks are of different kinds and require the learners to make guesses, based on their general knowledge and personal experience, during which they use their imagination and test their power of reasoning.

● **Discussions and decisions:** Require the learners to share their information to reach a decision.

Richard (2001, as cited in Nunan, 2003) provided the following five task types:

● **Jigsaw tasks:** In these kinds of task, every learner or a group of learners has a piece of information. They combine the pieces to form a whole.

● **Information gap tasks**: tasks in which one group of learners are given a set of information and another group are given a complementary set of information. They negotiate in order to understand the other group' information to compete an activity.

● **Problem solving tasks:** In these kinds of tasks learners are given a problem and a set of information. They must negotiate and find a solution to the problem, and there is generally a single outcome.

● **Decision making tasks**: In these kinds of activities learners are given a problem with a number of possible outcomes, and finally they select one through interaction and discussion.

● **Opinion exchange tasks**: In these kinds of task learners discuss something and exchange their opinions.

D. Accuracy, Fluency and Complexity

According to Skehan (1996), three aspects of language, namely, fluency, accuracy and complexity are in competition for attentional resources. Fluency refers to the learners' ability to produce language fluently in real time without hesitation (Skehan, 1996, as cited in Ellis, 2003). Learners can be fluent speakers through memorizing integrated language elements. Accuracy is the ability to avoid errors in performing the target language.

According to Skehan (1996 as cited in Willis & Willis, 1996), accuracy concerns "how well language is produced in relation to the rule system of the target language" (p.22). And complexity concerns the learners' ability to take risk and to expand their inter-languages. For example the number of clauses per T -unit or C-unit. And it is the ability to use more advanced language automatically. This aspect is correlated with restructuring, i.e. the development of interlanguage system.

Different researchers have different opinions about measuring accuracy. Some examine how accurately some grammatical features (like tenses) are used; others have chosen more generalized features or measures, such as percentage of error free clauses that don't contain any error. For instance, in the study done by Yuan and Ellis (2003) themselves, the same definition was made. They related all errors to syntax morphology and lexical choice.

III. RESEARCH METHOD

A. Participants

The participants of the present study were 60 females EFL learners studying in one of the language institutes (Sahand institute) in Miandoab. Their ages ranged from 14 to 18. The course offered for the classes was eight-units from Select Readings.

A pre-test consisting of 30 items grammatical judgment test was administered to 60 learners. The testees were divided into two groups according to their scores in pre-test. The instructor tried to have two linguistically homogeneous groups, then one of the groups consisting of 30 students was randomly selected as experimental group and the other one selected as control group. Both groups received the same amount of instruction (about 8 sessions), using the same material taught by the same instructor. Finally students in both control group and experimental group were administered the same post-test.

B. Instrumentation

Grammatical judgment test is taken to be one of the important instruments which was used in the present study. Before administering the grammatical judgment test the reliability and validity of the test were estimated. For estimating reliability the researcher used test-retest reliability. In test-retest reliability according to Hatch and Farhady (1981, p. 246), "Reliability is obtained by administering the test to the same students twice and computing the correlation between the two administrations". The researcher administered the test twice to 20 EFL learners and then computed the correlation between the two administrations. And also validity of the test was computed through content validity. Hatch and Farhady (1981, p. 250) defined content validity as "the extent to which a test measures a representative sample of the subject matter content". The focus of content validity is on the adequacy of the sample and not simply on the appearance of the test. To assure content validity of this test, the content of whatever the researcher wish to measure carefully defined. Eight reading passages from select readings were another instrument that was used in this study.

C. Procedures

Firstly, 60 students from one of the institutes (Sahand institute) were given a pre-test including 30 items grammatical judgment test adapted from Nelson English language tests (specially, tense and pronoun). The testees, then, were divided in two 30-member groups on the basis of their obtained scores. The instructor tried to put the same number of students who gained almost the same scores in both groups. Then, one of the groups was randomly chosen as experimental group to receive the treatment.

Second, researcher created small teams among the experimental and control groups according to principles of jigsaw task. Here, too, regarding the students' pre-test scores, the instructor tried to make six equal teams out of the 30 participants in the experimental group and make six equal teams out of the 30 participants in the control group. That is, each small group included two weak students who received below 18, two average students who received between 18-32, and one strong student who had got from 32-40 in the pre-test. Consequently, the experimental and control groups were divided into six teams, each with five members to work together for the purpose of retelling the part of the reading passage according to principles of jigsaw task. In these kinds of tasks, every learner or group of learners has a piece of information. They combine the pieces to form a whole (e.g. two or three groups of learners have different parts of the story and put the pieces together to complete the story) (Richards, 2001 as cited in Nunan, 2003). In the present study, according to principles of jigsaw task, first paragraph was given to group one for retelling, second paragraph was given to group two and third paragraph to group three and so on.

Then, the researcher corrected the learners' errors in the experimental group after they finished the retelling the passage (metalinguistic feedback). The instructor only corrected the errors related to a tense and pronoun. In each session, they covered one reading passage from Select Readings. Learners in control group didn't receive any feedback.

Both the experimental and control group lesson plans were based on the same reading selections. However, the experimental plans provided opportunities for receiving treatment through metalinguistic feedback. Conversely, students in the control group didn't receive any feedback.

Finally, Students in both control and experimental groups were administered the same post-test, grammatically judgment test.

The mean and standard deviation of both control and Experimental groups were computed. The data collected was computed through t-test.

D. Design

The research question proposed in the present study requires Quasi- experimental mehod of research. Accordingly, the study employed a pre-test and post-test, control group, experimental design while focusing on the variables of metalinguistic feedback as an independent variable and grammatical accuracy as a dependent variable which was hypothesized to be influenced by the independent variable.

E. Measures

To measure the grammatical accuracy, at the pre-test stage after gathering the learners' scores, the researcher computed the mean and standard deviation of both groups. To be confident that both groups were homogeneous and there was no significant difference among them before the treatment, their obtained mean scores were compared through t-test. At the post-test stage, to prove the efficiency of the given treatment, also the same test as a post-test was administered for both control and experimental groups to examine differences after 6 sessions.

After gathering both experimental and control groups' scores, the means, ranges, and standard deviation for both groups were computed. Then another t-test was run to check the significance of the difference between post-test means of groups.

IV. ANALYSIS RESULT

As already explained, the pre-test including 30 items grammatically judgment test (including tense and pronoun) was given to clarify the actual linguistic condition of samples in both groups before treatment. Before administering the grammatical judgment test the reliability and validity of the test were estimated. For estimating reliability the researcher used test-retest reliability. And also validity of the test was computed through content validity. To assure content validity of this test, the content of whatever the researcher wish to measure carefully defined. After administering the test, learners' scores were gathered the mean and standard deviations of both groups were computed. After treatment, which lasted for six sessions, post-test was given to both control and experimental groups to observe the difference resulted from the treatment. Again the mean of both groups was computed as well as their standard deviations. However, to be confident that both groups were homogeneous and there was no significant difference between them before the treatment, their obtained mean scores were compared through t-test. To be sure about the significance of the difference between post-test means of both groups, another t-test was run to check the significance of the difference between post-test means of groups.

V. CONCLUSIONS

The first research question in the present study was: Does metalinguistic feedback have any effects on grammatical accuracy of EFL learners? The findings indicated that the metalinguistic feedback has a positive effect on grammatical

accuracy of the learners. Simply, when the participants received feedback through metalinguistic comments in the experimental group, there were higher mean differences between the two groups in the post-test.

The second research question addressed the effect of metalinguistic feedback on accuracy of pronoun. The results indicated positive effect of metalinguistic feedback on accuracy of pronoun.

The third research question was formulated to explore the effect of metalinguistic feedback on the accuracy of tense. The results indicated that the effect of metalinguistic feedback on grammatical accuracy of tense is greater than pronoun.

APPENDIX

TABLE 4.1:
THE MEAN, STANDARD DEVIATION AND STD. ERROR MEAN IN PRE-TEST AND POST-TEST
Group Statistics

	Group	N	Mean	Std. Deviati	Std. Err Mean
Pre test	control	30	24.325	7.7306	1.2223
	experiment	30	23.650	7.0267	1.1126
Post- test	control	30	28.475	9.5218	1.5055
	experiment	30	32.900	9.7869	1.5474

Table4.1 is the comparison of the pre-test and post-test results of the experimental and control groups. Table 4.1 shows that the pre-test mean score was 24 for control group and 23 for experimental group in pre-test stage. As it is obvious, the means of both experimental and control groups on the pre-test stage were homogeneous, while at the post-test stage the mean of control group is 28 and for experimental group is 32. It is concluded that the difference between the means of experimental group and control groups in the post-test is noticeable.

TABLE 4.2:
T-TEST FOR EQUALITY OF MEANS IN THE PRE-TEST AND POST-TEST STAGES

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
PRE	Equal variances assumed	.426	.516	.408	78	.684
	Equal variances not assumed			.408	77.320	.684
POST	Equal variances assumed	.128	.721	-2.050	78	.044
	Equal variances not assumed			-2.050	77.941	.044

In this table mean difference, t value, degree of freedom and 2-tailed probably are shown. As it is obvious from Table 4.2 there was no significant differences between the means of two groups at the pre-test stage. (-t = .408, df = 78, P = .684 > .05). And we could claim that both groups to be equal before the treatment. While there does seem to be some differences between them on the post-test and this difference is statistically significant. (-t = 2.050, df = 78, P = .04 < .05).

TABLE 4.3:
THE MEAN, STD. DEVIATION IN PRE-TEST AND POST-TEST
Descriptive Statistics

	Group	Mean	Std. Deviation	N
PRE	control	24.3250	7.73068	30
	experiment	23.6500	7.03672	30
	Total	23.9875	7.35276	60
POST	control	28.4750	9.52187	30
	experiment	32.9000	9.78696	30
	Total	28.6875	9.84898	60

Table 4.3 shows the mean and standard deviation of both control and experimental groups at pre-test and post-test stages. As it is clear, the mean scores of two groups at pre-test stage is homogeneous but at the post-test stage there was difference between means of two groups.

TABLE 4.4:
THE MEAN, STANDARD DEVIATION AND STD. ERROR MEAN OF BOTH CONTROL AND EXPERIMENTAL GROUPS FOR TENSE

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
TENSE	control	30	13.1000	5.90437	.93356
	experiment	30	24.4750	5.64318	.89227

Table 4.4 explains the mean and standard deviation of control and experimental groups for tense. The mean score for control group was 13.1000 and 24 for experimental group. And the standard deviations were respectively 13 and 24 for the groups. As it is obvious, the experimental group has higher mean in comparison to the control group.

TABLE 4.5:
T-TEST FOR EQUALITY OF MEANS FOR TENSE

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
TENSE	Equal variances assumed	.067	.795	-5.711	78	.000
	Equal variances not assumed			-5.711	77.841	.000

In this Table mean difference, t-value, degree of freedom and 2-tailed probably are shown. As Table 4.5 shows there was significant difference between the means of two groups. (-t = 5.71, df = 78, P = 0 < .05).

TABLE 4.6:
THE MEAN AND STANDARD DEVIATION OF BOTH CONTROL AND EXPERIMENTAL GROUPS FOR PRONOUN

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
PRONUN	control	30	15.6750	7.14820	1.13023
	experiment	30	19.9760	5.78795	1.07327

Table 4.6 explains the mean and standard deviation of control and experimental groups for pronoun. The mean of control group is 15.6750 and for the experimental group it is 19.9760. As it is obvious the difference between the means of experimental and control groups for pronoun is noticeable.

TABLE 4.7:
T-TEST FOR EQUALITY OF MEANS FOR PRONOUN

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
PRONUN	Equal variances assumed	..032	..851	-2.759	78	.007
	Equal variances not assumed			-2.759	77.792	.007

Table 4.7 indicates the t-value, degree of freedom and 2-tailed probability of the two groups for pronoun. As it is obvious the difference is statistically significant. (-t = 2.75, df = 78, P = .007 < .05

REFERENCES

[1]   Ahamad Shah, M.I. (2003). Language learning content-based English as second language (ESL) Malaysian Classrooms. *Journal of Language and Learning 1*, No. 2, PP.1740 – 4983.

[2]   Brandet C. (2008). Integrating feedback and reflection in teacher preparation. ELT Journal 62(1), 37-46.

[3]   Ellis, R. (2003). Task-based language learning and teaching. Oxford:Oxford University Press.

[4]   Ellis, R. (2001). Investigating form-focused instruction. *Language learning, 51*(1), 1- 46.

[5]   Ellis, R. (2003). Becoming grammatical. *Language Teaching & Learning, 7, pp.* 124-232.

[6]   Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, *35*(3), 407-432.

[7]   Grami, M. A. (2005). The Effect of teachers' written feedback on ESL students' writing: A Study on a Saudi ESL University-Level Context *,Language Teaching*, 2, pp. 9-22.

[8]   Gulcat, z; & Ozagac, O. (2004). Correcting and giving feedback to writing. *Journal of Second Language Writing,* 7. Retrieved Dec.12, 2014. from https://www.buowl.boun.edu.tr/teachers.

[9]   Harmer, J. (2001). The practice of English language teaching. (3 rd ed.). London: Longman

[10]   Hatch, E. & Farhady, H. (1981). Research design and statistics for Applied linguistics. Tehran: Rahnoma Publication.

[11]   Johnson, K; & Johnson, H. (Eds). (1998). Encyclopedic dictionary of applied linguistics. Oxford: Blackwell Publishers.

[12]   Lyster, R; & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in second language acquisition***,** *Vol. 19*, PP. 37-66.

[13]   Naeini J. 2008. Error Correction: an indication of consciousness-raising. *Novitas Royal*, 2(2), 120-140.

[14]   Nunan, D. (1989). Designing tasks for the communicative classroom. Published in the United Kingdom at the University Press, Cambridge.

[15]   Nunan, D. (2003). Task based language teaching. Published in the United States of America by Cambridge University Press, New York.

[16]   Nunan, D. (1999). Second language teaching and learning. Boston. Heinle & Heinle Publishers.

[17]   Nunan, D. (2001). Aspects of task-based syllabus design. *Linguistic issues*. Retrieved July 12, 2007 from http: //www 3. telus.net/linguisticsissues/syllabusdesighn.html.

[18]   Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds); *Challenge and change in language teaching* (pp. 22). Macmillan Heinemann English Language Teaching.

[19]   Williams, J. (2003). Providing feedback on ESL students' written assignments. *The Internet TESL Journal, Vol. 7*, No. 10, Retrieved October, 2003 from: http://iteslj.org/Techniques/Williams-Feedback.html.

[20]   Willis, J; & Willis, D. (1996). Consciousness-raising activities. In J, Willis & D. Willis (EDS), *Challenge and change in language teaching* (pp 63-65). Macmillan Heinemann English Language Teaching.

[21]   Yuan, F; & Ellis, R. (2003). The effects of pre-task planning and online planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, pp 1-27.

Somaieh Abdollahzadeh was born in Miandoab, a city in the west of Iran in 1982. She is a PhD student at Azad university of Tabriz. She received her MA in English language teaching from Azad university of Tabriz in 2008, her BA in ELT from the same university in 2005. From 2007 to 2011 she was an English language teacher in language institutes in Miandoab.

Since 2006 to date she has been working as a teacher at Azad university of Miandoab and Payamenoor University. Her current research interests include scaffolding, focus on form, foreign language learning.