Investigating the Effects of Written Output and Input Enhancement on EFL Learners' Grammatical Development

Azadeh Baghban Ferdous [*]

Shahid Beheshti University, Tehran, Iran

ABSTRACT

This study investigated the relative effectiveness of two types of instructional treatments, input enhancement, and output treatment via text reconstruction activity for teaching five target forms, involving 60 Iranian learners of English. The target forms were selected to address the noticing issue and noticing/triggering function of output in language learning. By administering the KET test, the researcher made sure of the learners' homogeneity. Then, the participants were assigned to two conditions, input and input enhancement (non-output group), and input and output (the output group). After a brief grammar explanation, the learners in the non-output group engaged in comprehension of reading texts containing the target forms and underlined the target forms (as an input enhancement technique). The learners in the output group were given the same grammatical explanation and the same reading texts. However, they were given the opportunities to reconstruct the texts as accurately as possible. The participants took two kinds of tests, namely a recognition test and two production tests. The performance of each group was compared by analyzing the learners' scores. The results of data analysis indicated that the two treatment groups were not significantly different in terms of their acquisition of target forms in the recognition test. However, the findings showed the superior performance of the output group than the non-output group in the production tests. The results suggest that although input enhancement techniques have benefits for learners' linguistic development, the output treatment due to its reflective nature and higher cognitive demand can lead learners to higher linguistic development.

Keywords: language awareness, informed-learning, form-focused instruction, input and output hypothesis, noticing, task-based language learning, language reconstruction, psycholinguistics, deep learning, contemplative learning

INTRODUCTION

The teaching of grammar continues to occupy a major place in language pedagogy and there has always been the question of how to help learners achieve greater linguistic benefit in the classroom. Larsen-Freeman (1991) believes that " linguistic accuracy is as much a part of communicative competence as being able to get one's meaning across or

---

to communicate in a sociolinguistic appropriate manner" (p. 280). Thus, harmonious with the above assumption, Larsen-Freeman defines teaching grammar as enabling language learners to use linguistic forms accurately, meaningfully, and appropriately.

It has been noticed in the students' performance that they are unable to use their grammatical knowledge as used by native speakers. Having witnessed the production of artificial, pidgin-like utterances and doing poorly in classroom grammar tests, the researcher intends to deal with these weaknesses in learners' performance, through examining the effect of two methods of making learners aware of their linguistic problems on their performance.

The traditional production practice has been an issue due to its lack of naturalness and authenticity. In addition to this issue, as Izumi and Bigelow (2000) believe, among the many methods and techniques that aim to facilitate the development of the interlanguage grammar, the role of output in SLA is relatively unexplored. A common assumption is that output is only an indication of what has already been learned and does not play any role in learning new linguistic items. But since Swain's output hypothesis came to the fore, this assumption has come into question. Taking triggering/noticing function of output into account, this study intends to explore the role of output by using text reconstruction task as a technique in making learners aware of their linguistic problems in their production.

LITERATURE REVIEW

Language awareness

Language awareness has been conceptualized in several different ways. Donmal (85) defines language awareness as a person's sensitivity to and conscious awareness of the nature of language and its role in human life (VanLier, 2001). Language awareness assumes that some form or level of awareness about linguistic use, knowledge, and learning is beneficial for learners (VanLier, 2001, p.160). In order to understand the role of language awareness in second and foreign language teaching, we have to consider whether awareness, attention, and noticing particular features of language adds to learning.

Attention in SLA

The global consensus that has emerged from decades of research in SLA is that input plays a significant role in driving learners' acquisition of an L2 (Izumi, 2002, p.542). Current SLA research however goes beyond general interest in the need for comprehensible input (Krashen, 85), which is considered necessary but insufficient (seeEllis, 94; Larsen Freeman& Long, 91 for reviews). Instead it seeks to obtain a more precise understanding of how learners process or interact with input to develop their interlanguage competence.

A general finding of series of research in cognitive psychology indicates that not all of the input that learners are exposed to is utilized as intake for learning. Indeed,

researchers have examined the role of attention in mediating input and learning (Izumi, 2002).

Categories of noticing

An attempt has been made to categorize noticing based on what is noticed. Swain (1998, p.66) pointed out that there can be different types of noticing such as:

Noticing a form in the input: this type of noticing is when, while the learner listens to reads the target language input, the learner simply attends to the formal aspects of the target language in the input. Input frequency, saliency of the input, and the external manipulations such as input enhancement, may influence this first type of noticing.

Noticing one's interlanguage deficiencies (holes): it is when learners may notice that they cannot say what they want to say precisely in the target language. This type of noticing is also referred to as noticing 'holes' (Doughty & Williams, 1998, P.228; Swain, 1998,p.66).

Noticing the 'gap' between the interlanguage and the target language: it is when learners may notice that their current interlanguage is different from the target language. Feedback provided during interactions may help learners notice their errors, that is, notice the gap (Long, 1996).

Integrative processing

Graf (1994) explained integrative processing as follows:

"Integration focuses on connections among the units that define an individual item, such as word, an object, or a sentence; these kinds of connections are formed or strengthened when the subject either perceives coherence among separate stimulus components (e.g., under the guidance of preexisting representations or gestalt laws like proximity or common fate) or conceives a structure for processing target features as a single entity. (Cited by Izumi, 2002, p. 570)"

Izumi(2002) contends that the notion of integrative processing is of great relevance to SLA because it underscores the importance of not only attending to individual form elements that make up a structure but also of perceiving or conceiving the relationship among them to cause structure(grammar) learning. Izumi attributes output processing by the output tasks as the stimulator of integrative processing and as the glue to connect individual form elements, which, one might say, were only vaguely related to one another during the comprehension process.

Output Hypothesis

The emergence of FonF debate invited many pedagogical proposals, among which Swain's work on *Comprehensible Output* and *Pushed Output* (1995) is seminal:

"It is while attending to produce target language (vocally or subvocally) that learners may notice that they do not know how to say (or write) precisely the meaning they wish

to convey. In other words, under some circumstances, the activity of producing the target language may prompt second language learners to recognize consciously some of their linguistic problems." (Swain, 1998, p. 67; the noticing role of output)

She also claims that in producing pushed output, learners will on occasion be aware of (i.e., notice) a linguistic problem. Noticing a problem can push learners to modify their output. In doing so, learners may be forced to a more syntactic processing mode than might occur in comprehension. Thus, output sets "noticing" in train, triggering mental processes that lead to *"modified output."* (Swain & Lapkin, 1995, p. 371). Swain and Lapkin (1995), also, suggest that what goes on between the original output and its reprocessed form is part of the process of second language learning.

Since the Output Hypothesis was first proposed, Swain (1993, 1995, 1998) has refined her hypothesis and specified the following four functions of output: (Cited in Izumi, 2003, pp. 170-1)

Fluency function: in order to develop speedy access to extant L2 knowledge for fluent productive performance, learners need opportunities to use their knowledge in meaningful contexts, and this naturally requires output. That is, output provides opportunities for developing automaticity in language use.

Hypothesis-testing function: producing output is one way of testing one's hypotheses about the TL. Learners can judge the comprehensibility and linguistic well-formedness of their IL utterances against feedback obtained from their interlocutors.

Metalinguistic function: It is claimed, Izumi cites Swain(1995), that ' as learners reflect upon their TL use, their output serves a metalinguistic function, enabling them to control and internalize linguistic knowledge'. In other words, output processes enable learners not only to reveal their hypotheses, but also to reflect on them using language. Reflection on language may deepen the learners' awareness of forms, rules, and form-function relationships if the context of production is communicative in nature.

Noticing/triggering (or consciousness-raising) function: In producing the TL ' learners may notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know partially' (Swain 1995; cited in Izumi, 2003). The recognition of problems may then prompt the learners to attend to the relevant information in the input, which will trigger their IL development.

In sum, Izumi (2003) concludes that, Swain's output can under some conditions, promote language acquisition by allowing learners to try out and stretch their IL capabilities.

Task-based Language Learning

The Interaction Hypothesis suggests a number of ways in which interaction can contribute to language acquisition. In general terms, it posits that, the more opportunities for negotiation (meaning and content) there are, the more likely acquisition is. More specifically, according to Ellis (2003, p. 80), interaction suggests:

"That when interactional modifications lead to comprehensible input via the decomposition and segmenting of input, acquisition is facilitated. That when learners receive feedback, acquisition is facilitated. That when learners are pushed to reformulate their own utterances, acquisition is promoted."

These claims provide a basis for investigating tasks. Tasks that stimulate negotiation and through this provide comprehensible input and feedback and push learners to reformulate are the ones that will best work for acquisition. Also, from the Noticing Hypothesis view, according to Thornbury (1997), two kinds of noticing are necessary conditions for acquisition:

"Learners must attend to linguistic features of the input that they are exposed to, without which input cannot become intake. Learners must "notice the gap", i.e., make comparisons between their current state of their developing linguistic system, as realized in their output, and the target language system, available as input. (p.326)"

Two terms, similar in meaning have been proposed for this second type of noticing, that is to say, "matching" and "cognitive comparison".

In trying to find an answer to the question of what kinds of activities and procedures might be conducive to noticing and matching, Thornburry (1997) addresses two techniques that have noticing and matching as their bases; namely, reformulation and reconstruction tasks.

Reformulation activities

Thornburry (1997) defines reformulation as

"A technique in the development of students' writing skills: rather than simply correcting a students' composition, which usually involves attention to surface features of the text only, the teacher reformulates it, using he content the student has provided, but recasting it so that written draft approximated as closely as possible to a putative target language model. It is then available for the student's own draft. "(p. 327)

According to Thornburry (1997), the technique has also been promoted for the teaching of speaking skills on-to-one classes, by the content being dictated by the student, the form only be the teacher.

Reformulation, he contends, then, reverses the order of traditional models of instruction, which move from accuracy to fluency, rather, it is consistent with a fluency-to-accuracy, or task-based, model of instruction.

Reconstruction Activities

Unlike reformulation activities, following Thornburry (1997), in which the learner's text is reformulated by the teacher, the starting point for reconstruction activities is the teacher's text (or, at least, a text provided be the teacher which the learner first reds (or

listen to) and then reconstructs. The reconstruction version is then available for 'matching' with the original.

In reconstructing a text, learners will deploy their available linguistic competence, which is likely to fall short of the target model. In forcing attention to form, this process alone, Thornbury states, activates bottom-up processes that, in comprehension, and communicative activities, are not necessarily engaged. Two beneficial processes that may occur during reconstructing texts, according to Thornburry (1997), are:

Noticing: the activity of producing the target language may prompt learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2.

Matching: the comparison by learners of their version with the model provides them with positive evidence of yet-to-be-acquired language features, and this process of noticing, theoretically, converts input to intake, and serves to restructure the learner's developing linguistic competence. (p.330)

A reconstruction activity, that has been popularized is known as *dictogloss*, dicto-comp, or grammar dictation, the basic procedure of which is, the learner's listening to a short text once, or twice, and reconstructing it form memory. The reconstructed text is then compared with the original. (Thornburry, 1997)

RESEARCH QUESTIONS AND HYPOTHESES

This study focuses on answering the following two research questions:

- 1. Is there any significant difference between the output group and the non-output group in terms of their performance on the recognition test?
- 2. Is there any significant difference between the output group and the non-output group in terms of their performance on the two production tests?

Based on the proposed research questions, the following null hypotheses are formulated:

- $H_0 1$. There is no significant difference between the output group and the non-output group in terms of their performance in the recognition test.
- $H_0 2$. There is no significant difference between the output group and the non-output group in terms of their performance in the two production tests.

METHOD

Since randomization was not possible, the present study is quasi experimental, with the post-test only design, involving two experimental groups (experimental group 1 and experimental group 2). The experimental group1 (the output group) was the group who received text reconstruction as output treatment. Experimental group 2 (the non-output group) received the input, and input enhancement (underlining) treatment. The design scheme is represented below:

T1= Homogeneity test (Ket Test)

X1= Input enhancement (underlining) treatment

X2= Text reconstruction treatment (output treatment)

T2= Post-tests (recognition and production tests)

Thus,

Experimental group 1   ...   T1 X2 T2

Experimental group 2   ... T2 X1 T2

The independent variable in this study was noticing with two levels, and the dependent variable was accuracy in using the grammatical items again with two levels.

Participants

Sixty female students, with age range of 16 to 25, participated in the study. The study was carried out in Khorasan Language Institute where the researcher was teaching a communicative course. The participants were students who took the course to improve their general English proficiency and the researcher had the experience of teaching them the previous terms. The students were placed in the course by means of passing the previous course or by having had a placement test. After the administration of the Ket test, to make sure of the homogeneity of the participants who were sixty students attending four classes with 15 students each, the students were randomly assigned to two experimental conditions.

Instrumentation

The following research instruments were used in the present study:

KET (Cambridge Standard Key English Test)

As it wasn't possible for the researcher to do random sampling, Ket was used to ensure the homogeneity of the groups. It should be mentioned that all the students have passed the previous semester's final exam with the minimum %75 of the total score.

Recognition test

The instruments used as post-tests were a recognition test and two production tests. It is worth mentioning that all the items in both kinds of tests were constructed to measure the learners' ability on the five target forms intended for this study, and no item contained other linguistic forms.

The recognition test comprised of two sections, similar to TOEFL test, namely, Structure, and Written Expression sections. The test was constructed to measure the learner's performance on the intended five target forms. The first section, the Structure section, consisted of 25 multiple-choice items, and the second section, the Written

Expression section, consisted of 25 error identification items. The total score for this test was 50.

Production tests

Two separate production tests were used to measure the students' ability in producing well-formed sentences after receiving the treatment. The tests were as follows:

Transformation test: In this type of production test, the students were supposed to paraphrase the sentences, beginning each new sentence with the words given. They were asked to make any changes necessary but not to change the general meaning of the sentences.

The researcher included this test since it is suggested by Heaton (1988) that it helps to provide a balance when included in tests containing multiple-choice items. Also, Heaton (1988) believes that transformation item type is " the one objective item type which is closest to measuring some of the skills tested in composition writing, although transforming sentences is different form producing sentences" (p. 46)

In order to come up with students' consistent and pertinent answers to this production type of items, the researcher, as the test constructor, did her best to restrict the number of possible answers. This was checked in the pilot study phase of this research.

Completion Items Test: This type of production test was used as a post test, along with the recognition and transformation tests, to measure the students' ability to produce acceptable forms of language. This test, specifically, measures the ability to insert the most appropriate words in selected blanks in sentences. The words selected to be put in the blanks were supposed to measure the five target forms. The researcher, as the test constructor, tried to use context and related sentences in each part of the test in order to avoid ambiguity and maintain authenticity.

As the tests introduced above had been developed by the researcher, they were subjected to the validation process. The report of the validation process is presented next.

Forms in focus

The target forms were five grammatical points which were selected from the students' course book, New Interchange (2). The researcher singled out these target forms out of twelve grammatical forms which were to be covered in the course book.

Due to the structures' syntactic and semantic complexities, these target forms are assumed to be difficult for the students and as a result problematic in recognition and production. This was confirmed through internal observations of the researcher over the treatment period. The selected structures were as follows:

- The distinction between present perfect and present perfect continuous,
- Conditional types 1 and 2,

- Relative clauses,
- Passive sentences with by and without by-present and past,
- The distinction between the present participles and past participles used as adjectives.

Materials

The book used in this study was New Interchange (2). The students in the two groups, experimental groups 1 and 2, studied the same book and covered six units. The reading texts which were assigned to be given to the experimental group 1 as input for reconstruction, and to the experimental group 2 as input for comprehension and underlining were mostly taken from New Interchange (2) course book and work book. That is, out of ten reading texts, having been chosen for treatment, eight texts were from these books. Two of the reading texts were selected from two other sources, one from Intermediate Grammar in Use book, and the other was taken from the internet. The aim of the researcher in choosing the reading texts from these sources was their authenticity and appropriateness for the so called students' level. The texts were chosen on the premise that they were enjoyable to read, and linguistically valuable.

For each target form, two reading texts were chosen. The reading texts were in the form of short paragraphs (10 paragraphs), which did not exceed five sentences. In this way, the texts could afford the students' memory span and their handling and processing capacities.

Procedure

The study was planned to examine the use of input enhancement (underlining) treatment and output treatment on learners' recognition and production of the five target forms. Thus, as an experimental study, it compared two types of instructional treatment, in four classes (two experimental 1 and two experimental group 2 classes) of intermediate level to investigate the effect of written output and underlining technique on the acquisition of the target forms. Therefore, to accomplish this aim empirically certain procedures were followed.

For both groups, the classes were held three times a week, altogether 25 sessions, and each session took about one hour and forty five minutes. The plan of the treatment sessions, the same for both groups, is provided in the session schedules below:

- ➢ *Session 1 ……… administering the Ket test*
- ➢ *Session 4 ……… modeling session for the experimental group 1*
- ➢ *Sessions 5 & 7 ……… treatment on the Present Perfect Continuous and Present Perfect Simple (2 reading texts)*
- ➢ *Sessions 9 & 11 ……… treatment on Passive sentences (2 reading texts)*
- ➢ *Sessions 13 & 15 ……… treatment on Participles as adjectives (2 reading texts)*
- ➢ *Sessions 17 & 19 ……… treatment on Relative clauses as subject (2 reading texts)*
- ➢ *Sessions 21 & 23 ……… treatment on Conditional types 1 & 2 (2 reading texts)*
- ➢ *Sessions 24 & 25 ……… administering post tests (recognition and production tests)*

In the first session, as a first attempt, a standard Ket test (a test of general proficiency) was administered in order to ensure the homogeneity of the two groups. An independent t-test was run to compare the mean scores of the experimental and the control groups on the Ket test. The t- test analysis result is provided in Table 1 below.

Table 1. T-test analysis for the Ket test

Groups	n	M	S	df	t	Sig
G1	30	65.7667	6.73	58	0.2	.847
G2	30	64.4000	7.24	57.702	0.2	

On the basis of Table 1, the probability equals 0.84 which is more than the level of 0.5. Thus, it can be claimed that the two groups were homogeneous in terms of their proficiency prior to treatment.

Afterwards, the researcher, randomly, assigned the four classes into two groups of 30 students, to undergo the two types of treatment.

After having been proved to be homogeneous, without being informed about the research aim, the participants were exposed to two conditions: input, underlining technique, and output (experimental group 1 and experimental group 2). The following is a brief description of the two types of treatment:

Experimental group1: Due to the students' unfamiliarity with the sequence of activities of the reconstruction tasks, one session (session 4) was dedicated to modeling the task and delineating guidelines on how to approach the reconstructing phase.

The modeling session included the teacher (researcher)'s explaining and modeling the activity including taking notes on the key words of sentences in a text, collaborating in reproducing the text, and comparing the text produced with the original text.

In each treatment session, the students, first, were given brief explicit grammatical explanations of the target form, using the section called Grammar Focus in the course book. It is worth mentioning the Grammar Focus section is preceded and followed by the use of the sentences in context. After a brief explicit grammar instruction, the students were given a reading text covering one target form in each treatment session. The students read the text for comprehension, answered the teachers' comprehension questions on-line, and then in their second reading, they took notes on key words of the text. They were given a time limit of 3 to 5 minutes, depending on the length of the text. Then, the students were asked to hand in the text to the teacher, in order not to get back to it when reconstructing the text. At this time, the students entered the reconstruction phase of the treatment. They were supposed to construct (reproduce) the processed text collaboratively (in pairs) and as accurately as possible to the original text. The students were told to finish reconstructing under the time pressure of not more than ten minutes. During taking notes, and reconstructing phases, the teacher checked out the class pair by pair, reminding the students of not writing sentences instead of key words, and not trying to memorize the sentences word by word instead of considering the message in each sentence. She encouraged the students to talk and discuss the

content and forms when reconstructing the text. As the last phase, the researcher called on a student from a pair of students to write their reconstructed text on the left side of the board, and another student to write the original text on the right side of the board. Then, the two texts were compared with one another, to highlight the erroneous forms to be corrected according to the original text.

The Experimental group2: like experimental group 1, the students in experimental group 2 were given a brief explicit grammar instruction on the target form in each session. Then, they were given the text and were told to read for comprehension. After answering the teacher's comprehension questions, the students were asked to underline the intended grammatical forms. In this phase, the teacher made the students aware of the grammatical forms by discussing the form in the class with the students and encouraged them to underline the form while paying attention to its use in the context. The treatment session was called off when the students completed underlining the grammatical form.

RESULTS

In this section, the results of the analysis of test scores obtained from the post-tests which were administered after the treatment phase of the study are presented. Addressing the issue of acquisition, the tests results indicate the relative improvement of the two groups after two and a half months exclusive treatments for each group.

The posttests included a recognition test of 50 items and two written production tests, namely, 25 item Transformation test and 25 item Completion test. The examination of the test scores in the recognition test was done with the scale of 50, one point was given for each correct target sentence and zero for each incorrect sentence. The analysis of the written production tests results was done with the overall scores of the two production tests, that is to say, the scale of 50. The scoring procedure in the production tests was carried out by giving one point for each target form used correctly in the learners' production.

In order to address the two research questions in this study, the researcher collected the data from the post-test scores and analyzed them. Since there were two groups in the study and all the assumptions of the t-test were met, the t-test formula (independent samples t-test) with .05 probability level was used to analyze and interpret the data.

Two separate t-test analyses were run on the learners' recognition and production tests scores. The first t-test was applied on the test scores obtained from the recognition test. The second t-test was applied on the overall scores of the production tests. The overall results are shown in Figure 1.

The Learners' Performance in the Recognition Test

The first research question is concerned with whether there is a significant difference between the two groups in terms of their performance on the recognition test. In order to answer the first research question, the data obtained form the test scores are

described graphically in Figure1. It shows the difference between the mean of the recognition test scores in the two groups. The figure shows a slight difference between the two groups' performance.
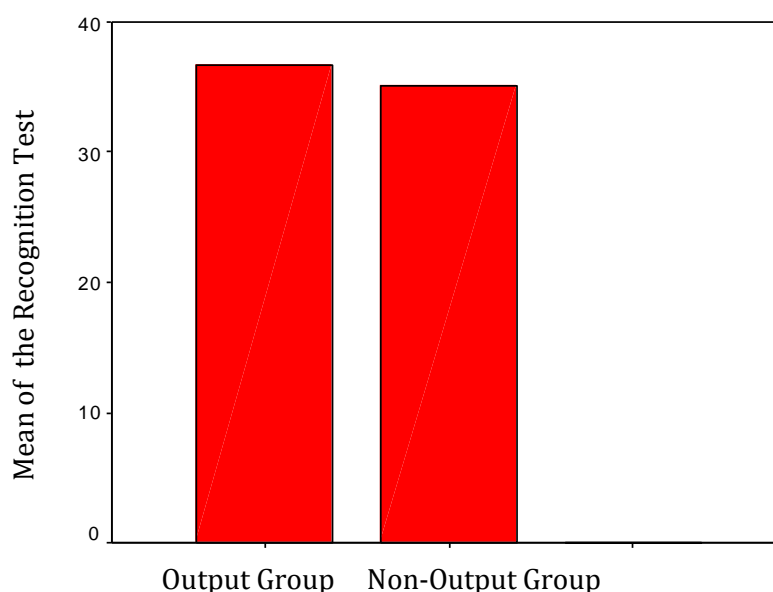

Figure 1. Means of recognition test scores of the two groups

In order to see whether the difference is significant, the researcher made statistical comparisons between the two groups, by submitting the recognition test scores to a t-test formula. According to Table 2, since the significance equals 0.4 and it is more than the α level of 0.5, the first null hypothesis cannot be rejected. In other words, the two groups showed no difference in linguistic development in the recognition test.

Table 2. Descriptive statistics of the recognition test scores of the two groups

Groups	n	M	S	df	T	Sig
G1	30	36.63	7.60	58	0.81	.42
G2	30	35.13	6.66	57.02	0.81	

The Learners' performance in the Written Production Tests

The second research question is concerned with whether there is a significant difference between the two groups in terms of the learners' performance on the production tests. In order to answer the second research question, the data obtained form the test scores are described graphically in Figure 2. It shows the difference between the mean of the production tests scores in the two groups. The figure shows a great difference between the two groups' performance.
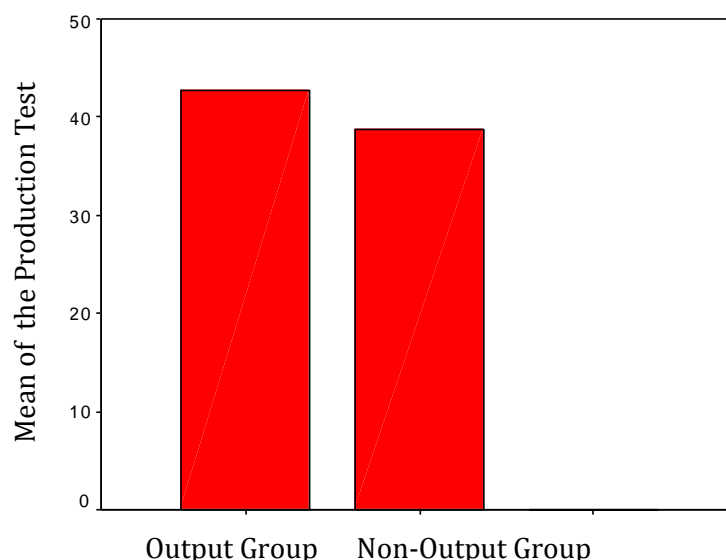
Figure 2. Means of production test scores of the two groups

In order to see whether the difference is significant, a t-test was run to answer the second question. According to Table 3, the significance equals .001 and it is less than the α level of .05. Therefore, the second null hypothesis is confirmed. In other words, there was a significant difference between the performances of the two groups with the output group performing better in the production tests.

Table 3. Descriptive statistics of the production test scores of the two groups

Groups	n	M	S	df	T	Sig
G1	30	42.63	3.63	58	3.88	0.001
G2	30	38.76	4.06	57.29	3.88	

The Comparison of the mean of the Groups in Recognition and Production Tests

In order to compare the results of the recognition test scores with the production test scores, the researcher provided the comparison of the mean of the output and non-output groups' test scores.

The figure 3 illustrates the degree of the groups' grammatical development, by comparing the means of the groups in the recognition and production tests. The figure consists of four columns. The first two columns show the non-output and output groups' mean scores, respectively, in the recognition test, indicating a slight superiority of the output group's performance in this test. However, as it was mentioned above, the t-test analysis inferred not a significant difference between the groups' grammatical accuracy in the recognition test.
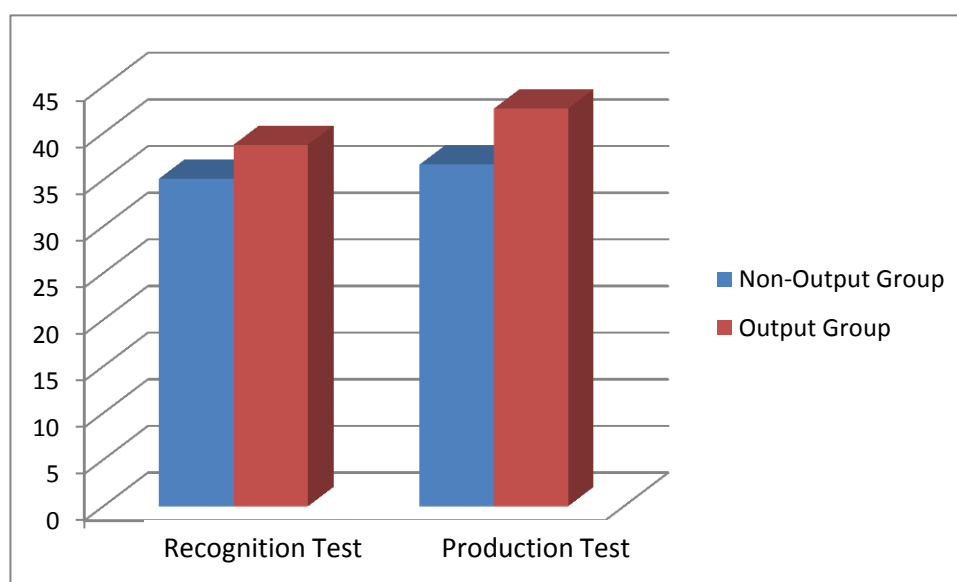
Figure 3. Comparison of the mean of the output and non-output groups in terms of recognition and production tests

The other two columns illustrate the non-output and output groups' mean scores, respectively, in the production tests, indicating the greater success of the output group in the production tests. Making a comparison between the recognition test columns (the first two columns) with the production tests columns (the second two columns), the figure provides the evidence that the output and non-output groups' grammatical development was undisputedly greater in the written production tests than in the recognition test.

DISCUSSION

To summarize the findings of this study in terms of the two research hypotheses, the results confirm the first null hypothesis, which predicted that there would be no difference between the two groups' linguistic ability in the recognition of the grammatical sentences. The results of the study support the idea that learners' noticing in the input and learners' noticing the 'holes' (mismatches) between input and their own output were equally effective in promoting grammatical accuracy of EFL learners in the recognition test. While the level of noticing in the non-output group was lower, compared to the output group, the non-output group did equally well in the recognition test as the output group. One possible reason which can be considered to account for this result is the nature of the recognition test. The recognition test, unlike the production tests, required less processing ability and cognitive demand on the part of the learners to answer the questions. In other words, since both groups had undergone noticing of the forms in the input and output, they might not have had problems at this level of test requirement, that is to say, recognizing the correct and incorrect grammatical forms. But the question that arises here is that why the output group did better in the production tests.

The second null hypothesis, which predicted that there would be no difference between the output and the non-output groups' linguistic ability on the production tests, was

rejected. The positive result of the output group in the production tests, as apposed to recognition test, can be an indication of the output group's deeper level of processing in the treatment phase of the study. Thus, on the basis of the findings we can claim that although underlining as a noticing technique, and output as noticing the 'holes' technique are internal priming devices, there would be a major difference between output and underlining. According to Izumi and Bigelow (2000), the difference lies in the fact that attention evidenced by the input enhancement (e.g., underlining) takes place at a relatively shallow level without necessarily shifting to deeper and more elaborate processing. In addition, output provides cognitive comparison which encourages further cognitive processing, compared to input enhancement, which may be necessary for acquisition. This argument is confirmed by the results of the present study.

In addition to the output learners' deeper level of processing, the notion of integrative processing is of relevance to this discussion. According to Izumi and Bigelow (2000), integrative processing underscores the importance of not only attending to individual form elements that make up a structure but also perceiving or conceiving the relationship among them to cause greater learning" (p. 571). In the present study, the output group not only attended to individual forms in the input but also focused their attention on how they are related to one another in the reconstruction and comparison phases of the output treatment, thus, having superior processing of grammatical forms in the input, compared to the non- output group, which had only the opportunity of attending to the forms in the input perceptually. Since the production tests, unlike the recognition test, required more integrative knowledge of the forms and more cognitive ability to produce sentences, the output group with that comprehensive treatment background showed greater performance in the production tests. This finding is consistent with Izumi and Bigelow and Izumi et al (1999)'s studies which explored this issue by comparing the performance of two groups of learners, under two conditions of output treatment and input comprehension and input enhancement. Izumi and Bigelow (2000) interpret the result of their study by claiming that input enhancement requires relatively shallow processing level, which led the learners to experience only a short-term retention of the attended form. On the other hand, greater learning evidenced by the output subjects suggests that output triggered deeper processing of the form, which led them to processing required in the output treatment, pushed the learners further in their cognitive processing and prompted them to perceive and conceive the utilized structure.

Although there was no control group to compare the improvement of the two experimental groups, the researcher would reject the idea that the non-output group did gain no grammatical accuracy after the treatment. This can be justified by the fact that since the output group showed development in this study, the equality of the results in the recognition test would be a sign that the non-output group had gained slight, if not high, grammatical accuracy; otherwise, the non-output group might have gained lower scores in the recognition test. Therefore, it is wise to conclude that underlining as an awareness raising technique was beneficial to some extent, or, at least

helpful for recognizing the grammatical sentences, but not helped for more comprehensive levels such as producing grammatical sentences. Thus, the researcher could relate this implied finding of the study to the studies which have seen awareness and attention to form as beneficial. One such study is that of Schmidt and Frota (1986), in which Schmidt analyzed his own acquisition of Portuguese during five-month stay in Brazil. Schmidt kept a diary of what he had noticed through instruction and also recorded his interactions with native speakers. By comparing the two sources of data, they found a significant association between recorded noticing in the form of diary entries and Schmidt's use of linguistic forms. As such, Schmidt cites this as "strong evidence of a close connection between noticing and emergence in later production" (1990, p. 141). Thus, according to the scholars such as Long (1998) and Ellis (1990), who have reviewed a number of empirical studies, we can conclude that conscious learning seems to contribute to successful language development.

However, concerning the use of underlining technique as noticing in the input, the researcher can interpret the result of the study as evidence for the lesser efficiency of underlining treatment compared with the output treatment. The finding can be justified in part by Sharwood Smith (1991, 1993)'s comments. She cautions researchers that it is inappropriate to assume that external manipulation of the input is the only mechanism that will increase learners' attention. Also, Doughty and Williams (1998) assert that input flooding and input enhancement may sometimes be too implicit to be maximally effective. They conclude that bringing the learner's attention to linguistic mismatch between interlanguage (IL) and target language (TL) is of the central issues in focus on form (FonF).

The findings in this study are consistent with a number of studies which have aimed to evaluate Output Hypothesis in general, and output as a tool for noticing/triggering (noticing the 'holes') in particular. The two significant studies which can be interpreted as support for the Output Hypothesis are Nobuyoshi and Ellis (1993) and Swain and Lapkin (1995)'s studies. Nobuyoshi and Ellis (1993)'s study support the Output Hypothesis in the sense that 'pushed output' focused on a specific linguistic aspect led to sustained improvement. Also, Swain and Lapkin (1995) reported on a study in which they looked at adolescent learners' awareness of gaps in their linguistic performance and the way in which the learners dealt with these gaps and applied various strategies to overcome the problems. The study's outcomes show that the learners did indeed become aware of the gaps and applied various strategies to overcome the problems.

The significant finding of this study can be summarized as the striking effect of output functioning as noticing/triggering or noticing the 'holes' by means of reconstruction activity. In other words, on the basis of the results of the production tests, the researcher can claim that output promotes noticing under certain circumstances (reconstructing the texts). Swain (1998) confirms the noticing function of output by asking learners to reflect on their production collaboratively by means of dictogloss. Her results are inferred from the language related episodes (LREs) taken form the collaborative dialogues made by the learners doing the dictogloss. The results suggest

rather forcefully that these language related episodes (LREs), during which learners reflect consciously on the language they are producing, may be a source of language learning. She concludes that increasing the frequency of LREs in pedagogical contexts through appropriate modeling, and through opportunities for use, might be useful in promoting language learning.

Of other studies which show the significance of output as noticing/triggering function, the series of studies done by Izumi and his colleagues (Izumi 2002;Izumi, 2000; Izumi & Bigelow,2000; Izumi, Bigelow, Fujiwara, and Fearnow, 1999) are consistent with the findings in the current study. With regard to the specific role of learner output in noticing and L2 learning, Izumi (2002) argue that his study revealed a threefold facilitative effect of pushed output on SLA:

   a) It promoted detection of formal elements in the input;
   b) It promoted integrative processing of the target structure; and
   c) It promoted noticing of the mismatches between the learner's interlanguage form and the target language input.

Izumi concluded that the findings of his study showed that "pushed output can induce learners to process the input effectively for their greater IL development."

The collaborative nature of the reconstruction activities, during which the output group's learners talked about and discussed the content and form of the texts in pairs, would be another reason for the success of the output group in this study. The learners' communication when using language to co-construct the language they need to express the meaning they want and to co construct knowledge about language, served as a tool for L2 learning. Some studies like Swain (1998)'study (as mentioned in the preceding paragraph) supports our findings in this regard. One of the studies is LaPierre's (1994) study (see Swain &Lapkin, 1998) which investigated the grammatical development of grade 8 early immersion students. The task that students engaged in was a story reconstruction task. The result of LaPierre's study suggests that the language related episodes, where students reflect consciously on the language they are producing, provide opportunities forL2 learning. Also, Swain & Lapkin (1998)'s study provides empirical data to suggest that collaborative dialogue is a useful concept for understanding L2 learning.

In summary, based on the above discussion, this study suggests that output under some conditions, would promote learners' interlanguage. That is, learners as a result of producing the target language, working collaboratively to co construct the language, and making comparisons between their interlanguage (IL) and the target language (TL) would develop their grammatical accuracy and thus their interlanguage.

CONCLUSIONS

In summary, this study suggests that providing opportunities for learners to compare their own production to the processed input would promote language learning. In other

words, output forces learners to move from semantic to more syntactic processing, and thus, leads to the acquisition of the forms focused on in this study.

The findings in this study indicated two points:

1. The output-based instruction aimed at making learners aware of the mismatch between their IL and the TL is more effective for producing target-like utterances.

2. Although learners noticed grammatical forms in the input in the non-output group, they had difficulty producing target-like utterances.

We conclude from the findings that some form of attention is beneficial for learning to take place. Thus, without denying the essential role of input in language learning, this study leads us to incorporating activities which tend to promote learners' noticing by providing opportunities to compare their output to the original input.

REFERENCES

Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language learning*, *46*(3), 529-555.

Doughty, C., & Williams, J. (eds) (1998a). *Focus on form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (1998b). Issues and terminology. In C. Doughty and J. Williams (eds), *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (1998c). Pedagogical choices in focus on form. In C. Doughty and J. Williams (eds), *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.

Ellis, R. (1993). Second language acquisition and the structural syllabus. *TESOL Quarterly*, *27*, 91-113.

Ellis, R. (1994). *The study of Second Language Acquisition*. Oxford: Oxford University Press.

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

Farhadi, H., Ja'farpur, A.J., Birjandi, P. (1994*). Testing language skills form theory to practice.* Tehran: SAMT.

Hatch, E., & Farhadi, E. (1981). *Research design and statistics.* Tehran: Rahnama Publications.

Heaton, J.B., (1988). *Writing English language tests.* New York: Longman.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24,* 541-577*.*

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics, 24*, 168-196.

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: effects of output on noticing and second language acquisition, *Language Learning*, *21,* 421-452.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *Tesol Quarterly, 34,* 239-278.

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *The Canadian Language Review, 60*(5),587-609.

Larsen-Freeman, D. (1991). Teaching Grammar. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (pp. 279-296). New York: Heinle & Heinle.

Long, M. H. & Robinson, P. (1998). Focus on form: theory, research and practice. In C. Doughty and J. Williams (eds), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Nobuyoshi, J., Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*(3), 203-210.

Schmidt, R.W., (1990). The role of consciousness in second language learning. *Applied linguistics,11*(2), 129-158.

Schmidt, R.W., (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.

Schmidt, R. (1995). *Attention and Awareness in Foreign Language Learning*. Manoa: University of Hawai'i.

Smith, M. S. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second language research*, *7*(2), 118-132.

Smith, S. (1993). MA Input Enhancement in Instructed SLA: Theoretical Bases. *SSLA*, *15*(2), 165-179.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian modern language review*, *50*(1), 158-164.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty and J. Williams (eds), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: a step towards second language learning. *Applied Linguistics, 16*(3), 371-391.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320-337.

Thornbury, S. (1997). Reformulation and reconstruction: tasks that promote 'noticing'. *ELT journal*, *51*(4), 326-335.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in second language acquisition*, *16*(02), 183-203.

Van Lier, L. (2001). Language awareness. In Carter, R., & Nunan, D. (eds), *Teaching English to speakers of other languages*, (pp. 34-42). UK: Cambridge University Press.

Vickers, C. H., & Ene, E. (2006). Grammatical accuracy and learner autonomy in advanced writing. *ELT journal*, *60*(2), 109-116.

Williams, J. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition, 21*, 1-48.