

Output Tasks, Noticing, and Learning: Teaching English Past Tense to Iranian EFL Students

Mohammad Khatib

Department of English Language and Literature, Allameh Tabatabai University

North Allameh St., Sa'adat Abad, Tehran, Iran

Tel: 98-912-310-6816 E-mail: khatib@atu.ac.ir

Mehrasa Alizadeh (Corresponding author)

Department of English Language and Literature, Allameh Tabatabai University

North Allameh St., Sa'adat Abad, Tehran, Iran

Tel: 98-912-438-8474 E-mail: mehrasa.alizadeh@gmail.com

Received: January 6, 2012

Accepted: February 22, 2012

Published: April 1, 2012

doi:10.5539/elt.v5n4p173

URL: <http://dx.doi.org/10.5539/elt.v5n4p173>

Abstract

EFL Learners often have problems using the past tense accurately. In an attempt to solve their problem, this study was carried out to examine the effects of using two different types of output tasks on noticing and learning the English past tense. Sixty female school-age EFL learners were divided into groups of 18, 19, and 23 participants. A pretest was administered at the outset of the study, the scores of which proved that all the participants equally lacked the required accuracy in using the target structure. Therefore, five treatment sessions followed, during which the first two groups were given picture-cued writing tasks and reconstruction tasks respectively. The comparison group, however, did comprehension check-up tasks. Finally, a posttest was given. The results of the statistical analyses revealed that only the reconstruction group improved in their noticing of the target feature. However, both experimental groups equally promoted their learning of the form.

Keywords: Output hypothesis, Noticing, Past tense, Picture-cued writing task, Reconstruction task

1. Introduction

Mastering the grammar of a second language and being able to correctly implement this knowledge is a challenging task to accomplish. That is why most ESL/EFL learners often have problems using language forms accurately in oral and written production. They may have a good knowledge base of the L2 structures but might find it difficult to implement their declarative knowledge when it comes to practice. This very deficiency is what makes grammar instruction open to research (Ellis, 2006).

As Ellis (2005) states, research into naturalistic L2 acquisition proved that learners, more or less, followed the same sequence in acquiring grammatical structures. This led to the belief that they had their own built-in syllabus for learning grammar. Later, Krashen, in his famous input hypothesis, argued that language is acquired through receiving *comprehensible input*, moving from i to $i+1$ by understanding input that contains $i+1$ (Krashen, 1981, 1982, 1985, 2009; Krashen & Terrell, 1983; McLaughlin, 1987). According to him, grammar instruction could not at all contribute to the development of implicit, automatized knowledge, but could only lead to gaining explicitly learned knowledge. In other words, one should adopt a *zero grammar approach* in L2 teaching (Ellis, 2005). This belief was highly popular during Krashen's era.

It was in the 1980s that Swain first proposed the *output hypothesis* in response to Krashen's *input hypothesis*, based on her observation of French immersion programs in Canada, where she found the students to be much weaker in their oral and written production compared to their reading and listening comprehension abilities. She advocated more opportunities for learners to engage in language production (i.e. output) in order to promote their linguistic abilities (Lightbown & Spada, 2006; Swain, 1985, 1995; Swain & Lapkin, 1995).

Apart from promoting automaticity in language use, Swain (1995, 2005) outlined three major functions of output in L2 acquisition: (1) the hypothesis-testing function, (2) the metalinguistic (reflective) function, and (3) the

noticing/triggering function. Based on the preliminary function of output, the more learners produce verbal output, the better they will get at producing it automatically, without having many pauses and time lapses to think about what they want to utter.

The first major function has to do with the trial-and-error process in language learning. As a learner, one forms various hypotheses in his/her mind about the way a target language works. Output production, therefore, presents learners with opportunities to test the hypotheses they have already formed about the second language and see how they really work in practice.

Output production also aids learners in reflecting metalinguistically upon their own efficiency in using a target language, which, by itself, is effective in promoting their subsequent language use.

The last function, i.e. the noticing function of output, which is the focus of this study, is based on Schmidt's *noticing hypothesis* (Schmidt, 1983, 1990, 2001, 2010; Schmidt & Frota, 1986). Put in simple words, being conscious of the accuracy and appropriacy of one's utterances is impossible unless one starts talking.

Put together, it seems that teachers are not supposed to leave their learners alone in their struggle of learning a target language. There are ways through which they can facilitate for their learners what seems at first to be too difficult, and they can pave their way toward accurate and automatic language use.

In the following section, studies dealing with output production and language learning and teaching will be reviewed.

2. Previous Studies on Output

Swain and Lapkin (1995) were among the first to explore the cognitive processes generated by producing output. They attempted to determine if learners' output would serve as an attention-grabbing device, and if it did, whether it would stimulate learners to engage in linguistic analysis. The researchers worked with 9 adolescent students in a French immersion program in Canada. The participants were given an environmental topic to write about while they were thinking aloud. The think-aloud protocols were analyzed in terms of the language-related episodes (LREs) identified in them. Having classified the LREs into seven descriptive categories, Swain and Lapkin found that their young learners did become aware of the gaps in their linguistic knowledge as they produced a short text in French. Moreover, upon encountering difficulties in producing the target language, the learners engaged in thought processes that might have played a role in second language learning. In short, the study suggested that the need to communicate created by the writing task forced learners to think about the form of their linguistic output. That is to say, it pushed them to move from semantic to syntactic processing.

A few years later, Izumi and his colleagues conducted a series of studies to test the output hypothesis and in particular its noticing function. The first of these studies was Izumi, Bigelow, Fujiwara, and Fearnow (1999). These four researchers wished to improve Swain and Lapkin's (1995) study by addressing this question: whether the awareness of the shortcomings in one's interlanguage during output production could prompt the learner to look for subsequent relevant input more attentively. Against what was hypothesized, both the output and the non-output participants equally improved in their noticing of the target form. However, the output group partially outperformed the other one in its immediate incorporation of the target form. The third hypothesis, concerning greater gains in the accuracy of the use of the target structure among the output participants, was also partially confirmed. Overall, Izumi et al.'s (1999) study lent partial support to the output hypothesis. But its findings were not comprehensive due to some methodological challenges and low generalizability.

It could be that the findings of this study had been contaminated due to the order of the presentation of the tasks. Izumi and Bigelow (2000), thus, set out to explore the role of task ordering in noticing and learning.

Aiming to make sure about the comparability of the two studies, Izumi and Bigelow (2000) kept all the parameters the same as the preceding study except that they reversed the order of the tasks given. The results revealed that the output group's noticing of the target form was similar to that of the comparison group. Yet, the output group showed significant gains in its accurate use of the target form. Thus, it was more successful than the non-output group in its immediate uptake of the form. However, both groups improved significantly in their use of the form on the posttests.

Putting the findings of these two studies beside one another, it can be concluded that the treatment effect was found to be of a "cumulative nature". This means that extended opportunities to produce output and to receive appropriate input play a major role in enhancing learners' accuracy in using a given grammatical form.

In the same decade, Izumi (2000, 2002) examined the potential differences that output production vs. input enhancement would make upon noticing and learning the target structure, among adult ESL learners. He came up with four major findings after carrying out the experiment. First, it was found that those engaged in both input and

output activities outperformed those who were exposed to the input solely for the purpose of comprehension. Second, unlike the observed positive effect of output on learning the target structure, visual input enhancement failed to show any significant effect on learning. Third, the hypothesis that the effect of output on noticing and learning would be comparable to that of input enhancement was rejected. Finally, the study revealed a rather weak relationship between noticing of the form, as measured by note-taking and reconstruction scores, and learning of the form, as measured by the pretest and posttest scores.

Further research on the same issue was pursued in Izumi and Izumi (2004). They attempted to investigate the effects of oral output tasks on the learning of relative clauses in English. The output group, was first presented with input containing the target grammatical form and then asked to perform an oral picture description task. The non-output group, however, was also exposed to the same input, but they were later asked to do a picture sequencing task, which required no oral output. The comparison group received a placebo. Despite the researchers' expectations, the non-output group outperformed the other two groups, showing greater overall gains in learning the target form. Izumi and Izumi (2004) discussed their findings in terms of the cognitive processes involved in each of the treatments.

Apart from the above pioneering studies, there are other studies which were carried out in various ESL/EFL settings. An example in case is Hanaoka (2007), who was interested in examining the effect(s) of noticing on the performance of Japanese learners of English. Working in an EFL setting, he gave his participants a four-stage writing task, which consisted of written output production, comparison against a model text, and two revisions. The researcher found that learners were more attentive to lexical features, identified their problems, and made use of the model text to overcome their problems in subsequent revisions. Also, the results proved that more proficient learners outperformed less proficient ones in noticing more features in the input. Moreover, those non-target-like features, which were problematic to learners in their output, were better noticed and revised in their subsequent output than those features which were unrelated.

Mennim (2007) reported a study carried out among Japanese EFL learners at the tertiary level. The researcher intended to explore the long-term effects of noticing on learners' oral output by giving them classroom exercises that promoted noticing and conscious attention to form. During the course of an academic year, the students' records of noticing was followed, and their oral output was recorded and analyzed to check for potential signs of improvement in the use of the forms they had noticed. Specifically focusing on data from two students who had problems using a count noun, he observed improvements in their use of the form after nine months, thus compiling evidence in favor of the facilitating role of noticing in developing grammatical accuracy.

In another study, Song and Suh (2008) tried to investigate the effects of two types of output tasks (reconstruction tasks vs. picture-cued writing tasks) on noticing and learning of a grammatical feature, the English past counterfactual conditional, in comparison to a non-output task (reading comprehension). After a one-month treatment, a posttest, consisting of recognition and production tasks, was administered to the students to check for their noticing and learning of the past counterfactual conditional. In terms of noticing, it was found that output tasks indeed made a difference and were able to promote greater noticing of the target form compared to non-output tasks. Also, the output groups outperformed the comparison group on the production task, but no effect was found for output task type. That is to say, reconstruction and picture-cued writing tasks were equally effective in promoting the participants' learning of the grammatical form under investigation.

Applying minor changes to Hanaoka's (2007) methodology, Kong (2010) examined the role of written output in promoting learners' noticing and accurate use of problematic grammatical forms. She specifically focused on note-taking as a way of facilitating noticing. Having divided her participants into two groups, she gave them all a three-stage writing task, in which they were asked to write descriptions of a given picture and to compare their descriptions to a model text. The note-taking group was also allowed to jot down notes during the comparison stage, whereas the non-note-taking group was not allowed to do so. Following that, the participants were required to write their second versions of the descriptive text, one group by referring to their notes and the other not using anything. A retrospective questionnaire immediately followed the writing task in order to further delve into the "noticing process". The results indicated positive effects for note-taking in facilitating noticing and thus enhancing grammatical accuracy.

Although not extensively studied, output has also been exploited in language teaching and learning in Iran. Four research studies are reviewed here.

Ahmadi (2003) found positive effects of dictogloss tasks on his learners' performance on a cloze test and a recognition test as far as their knowledge of English participles was concerned. He read aloud several texts to them and asked them to collaboratively reconstruct those texts. The results proved that this type of output tasks can

increase learners' awareness of grammatical features and can also motivate them to use language to interact with their peers.

Tabatabai and Yakhbi (2010) were interested in comparing the effects of receiving input vs. producing output among female Iranian university students as far as grammatical accuracy and complexity were concerned. Having assigned one group to input tasks and the other to output tasks, they found out that the output group outperformed both the input and the comparison groups in that they produced more grammatically accurate and complex structures in their monologues. The researchers confirmed Swain's (1995, 2005) claims that output production would lead to promoted language learning.

Working with male young-adult learners, Abadikhah and Zarrabi (2011) also investigated the role of output production on acquiring verbal morphology. They assigned a reconstruction task to their output group, which made use of an audio-text cartoon strip. According to the pretest scores, there existed a statistically significant difference between the participants' abilities in comprehending and producing verbal morphemes prior to the treatment sessions. However, the posttest scores revealed benefits for the output group over the comparison group in filling the gap between their comprehension and production of verbal morphemes. Therefore, the researchers found further support for the use of output tasks in L2 teaching and learning.

Replicating the study conducted by Song and Suh (2008), Ghari and Moinzadeh (2011) investigated the effects of using two different output tasks (picture-cued writing tasks vs. reconstruction tasks) on noticing and learning English past modals. Working with 94 adult Iranian learners, the researchers placed them in three groups, two experimental groups and one comparison group. At the beginning of the study, a pretest, consisting of recognition and written production tests, was administered to all the participants. Three treatment sessions followed, during which all the participants had to read a text containing the target form. After that, one of the experimental groups reconstructed the original text, given a general picture and some cues to lower the memory load of the task. The other group also did a guided writing task based on the pictures and the cues given. Finally, the comparison group answered comprehension questions. One week after the last treatment session, the posttest was given. Results of the comparisons between the pretest and posttest scores revealed that the two experimental groups outperformed the comparison group both in terms of noticing and learning the past modals. However, unlike Song and Suh (2008), Ghari and Moinzadeh (2011) found reconstruction output tasks to be more effective in promoting noticing of the target structure, but not learning it.

Glancing at the literature particularly from a localized perspective, it seems that the use of output tasks in language teaching and learning has not been well established in Iran. More specifically, various output tasks have not yet been experimented with among different age groups and with various target structures. Therefore, the current study attempts to test the effects of using two different kinds of output tasks upon learners' noticing and learning of the past tense in English. The researcher aims to answer the following research questions:

1. Do output tasks (picture-cued writing tasks and reconstruction tasks) make any difference in Iranian learners' noticing of the past tense compared to a non-output task (reading comprehension check-up tasks)?
2. Do output tasks make any difference in Iranian learners' learning of the past tense compared to a non-output task?

3. Methodology

3.1 Participants

The participants were Iranian female high school students studying at grade 2 in Farzanegan II High School in the 1st district in Tehran. They were all sixteen years old. These students had already had at least four years of training in English throughout their schooling. They were also familiar with three different aspects of the past tense (preterit, progressive, and perfect) since they had been taught these forms before as part of the curriculum developed by the Ministry of Education. In other words, the participants already knew about the past tense, yet the treatment given by the researcher was an attempt to consolidate their knowledge and contribute to their accurate use of the target structure.

The second graders were placed in three classes. These three classes represented the three groups of this study. In order to make sure about the homogeneity of the groups, a pretest was given to all the participants on the sixth session. Hopefully, the results of the pretest showed that there was no statistically significant difference among the groups as far as their knowledge of the past tense was concerned. The two experimental groups, from now on EG1 (N=18) and EG2 (N=19), were given the treatment for a period of 5 sessions, while the remaining 23 comparison students, from now on CG (N=23), were given a placebo.

EG1 was given picture-cued writing tasks while EG2 were presented with reconstruction tasks. CG also completed comprehension checkup tasks with no requirements of output production. The tasks given to each group and the

materials used are explicated in the following section.

3.2 Instrumentation

3.2.1 Pretest

Prior to the treatment sessions, a pretest was given to the participants, which included recognition and production items. The recognition items were comprised of 20 multiple-choice questions examining how far the participants would be able to recognize the appropriate choice from among four given alternatives. This section was adopted from *Objective tests in English as a foreign language* by Bloor, Bloor, Forrest, Laird, and Relton (1970).

The pretest also consisted of a part dealing with the production of the target form. The participants were given a simple narrative prompt to write a short text about a familiar topic. The prompt required them to use the past tense in accounting about an event which had already taken place, like writing about a vacation. They were given the topic sentence along with several verbs in the form of infinitives to use in their writing. The students were obliged to use those cues, but they were also free to use any other verbs they found necessary. The given cues served the purpose of ensuring a minimum number of verbs appearing in each student's text. Pictorial cues were avoided as they were more or less similar to the tasks that EG1 was supposed to go through during the treatment. This would put the first experimental group at an advantage compared to the other two groups and would, thus, contaminate the findings.

Each participant was scored based on the number of instances in which she had used the past tense correctly in her paragraph divided by the whole number of instances of past-tense verbs. That is, if a participant used the past tense 10 times in her paragraph, two of which were incorrect, she would get a score of 8/10 or 80 on a scale of 100.

3.2.2 Treatment Materials

The materials used during the treatment sessions differed across the three groups. All the participants, no matter what group they belonged to, were first exposed to reading texts containing the target structure. The reading passages were mainly of the narrative genre, accounting for past events. These texts (adopted from Hill, 1980) served as a source of input as they contained various forms of the past tense. While reading, the participants were asked to underline words or phrases which seemed important to them. Underlining, later, served as a measure of noticing. After reading a given text, each group was given a different follow-up task. EG1 was presented with picture-cued writing tasks. That is, a set of pictures along with some written cues were given to them, and they had to rewrite the story they had just read, based on the pictures and the cues.

EG2 was given a reconstruction task in the form of a dictogloss (or grammar dictation). A dictogloss is different from a dictation or spelling quiz in that it does not merely check students' orthography but also checks their knowledge of morphology and syntax. For this group, the researcher read the text two times at normal speed. EG2 was instructed to listen to the text carefully for the first time. Later, as the text was being read to them for a second time, they were told to take notes in the form of single words or phrases. After the second round of listening, EG2 participants had to rewrite the text as closely to the original text as possible both in terms of grammar and content. Refer to the appendix for a sample picture cued writing tasks and reconstruction task.

Finally, CG answered comprehension questions in the form of multiple-choice items, which avoided focusing on the target feature. Obviously, this group was not required to produce any output.

3.2.3 Posttest

The posttest was similar to the pretest in terms of format and level of difficulty. However, in order to avoid contaminating the data due to the practice effect from the pretest to the posttest, the items were modified, meaning that the structures were kept the same, but the context and the vocabulary were changed.

3.2.4 Measures of Noticing

Due to the nature of the construct "noticing", it was impossible to measure it in a single one-shot test (Schmidt, 1990, 2001). Therefore, noticing was measured in terms of the participants' underlines of the reading texts used during the input-flooding stage.

Underlines were scored as a fraction of the relevant underlines out of the whole number of underlined words or phrases. Underlines were considered relevant if they were among one of these categories related to the past tense: *was/were* & *-ing* (for past progressive), regular (*-ed*) & irregular past tense verbs (for simple past), *had* & *past participles* (for past perfect), and adverbs of time referring to the past. For example, if a participant underlined 10 words or phrases in a text, among which 8 belonged to one of the above categories, her score of noticing would be 80 on a scale of 100. In this way, a numerical measure of noticing was attributed to each participant on each treatment session.

Noticing is triggered once learners attempt to produce output and are faced with problems in expressing themselves (Swain, 1995, 2005). Such noticing of the gap between their own interlanguage and the target language model would push them to attend to relevant input more consciously the next time they are exposed to it. Taking all this into account, the participants of this study were only evaluated in terms of their underlines from the second treatment session onward.

3.3 Data Collection Procedures

At the beginning of the 6th session of the 2011-2012 school year, the pretest was given to all the 60 second graders at Farzanegan II High School. They were placed in 3 groups (two experimental and one comparison). A week later, the treatment started. The participants in all the three groups were initially given a reading passage containing the target structure, i.e. the past tense. Having read the passage once, they were required to read the text again and underline any word or phrase (but not complete sentences) that seemed important to them.

After this input-flooding stage, each group went through a particular type of follow-up task. EG1 was presented with a picture-cued writing task. That is, the participants in this group were asked to write a short text describing the event that had taken place in a set of pictures using the clues provided. Upon accomplishing the task, the participants in this group were given the original text, based on which the pictures were drawn. They were told to compare their own writing with the model text.

EG2 was given a reconstruction task, in this case a dictogloss. The text they were exposed to as the input was read to this group two times at a normal speed. For the first time, the participants were instructed to listen carefully and try to comprehend the text. Next, during the second round of listening, they were given pieces of paper on which they would be able to take notes. The researcher had told them in advance that they could not write complete sentences as notes, but only words or phrases. After that, using their notes, they were asked to reconstruct the text as closely as possible to the original one, both in terms of content and grammar. Having gone through the reconstruction task, they were given the original text so that they would compare it to their own reconstructed version.

Finally, CG was given comprehension questions, without being required to produce any output. The same procedure was repeated for the next four sessions. One week after the fifth treatment session, the posttest was administered. Figure 1 is a schematic representation of what the three groups went through during a single treatment session.

4. Results of the Statistical Analyses

To answer the first research question, which had to do with learners' noticing of the past tense, the three groups' mean scores on the noticing measure were compared against one another using mixed ANOVA. In response to the second research question, two MANOVAs were run. The alpha level for all the tests was set at 0.05 ($p < 0.05$).

4.1 Results of the Mixed ANOVA

Based on the number of relevant underlines in the reading texts, every single participant was assigned four separate noticing scores from the second treatment session onward. Table 1 presents the descriptive statistics for all the three groups on each of the four treatment sessions. It is worth reminding that the data on noticing collected during the first treatment session was discarded since it was not in line with the definition of noticing as proposed by Swain (1995, 2005).

With *time* (i.e. sessions) as the within-subjects factor and *group* as the between-subjects factor, a mixed ANOVA was run to assess whether there were group and time differences in the participants' noticing of the target feature. The assumptions of this test were met. Results indicated a significant main effect of time, $F(2.21, 126.00) = 6.47, p < 0.05$, and also of group, $F(2, 57) = 12.92, p < 0.05$. Furthermore, the time main effect was qualified by a significant interaction between time and group, $F(4.42, 126.00) = 2.91, p < 0.05$. This indicates that the three groups differed significantly in their noticing scores during each session, and that each group's (or maybe one or two group's) noticing scores differed significantly across sessions.

Tables 2 and 3 depict the details of the statistical analyses. The first table presents the results of the tests of within-subjects effects. In the first row, we see a significant effect of time as the within-subjects factor, with an η^2 of 0.10, which is a rather small effect size, according to Leech, Barret, and Morgan (2005). The second row shows the interaction between the within-subjects and the between-subjects factors (i.e. time * group). According to the figures in this row, the interaction between these two factors is also a significant one. Yet the effect size is again small ($\eta^2 = 0.09$), which means that only 9% of the variability in the data could be attributed to the interaction between time and group. Table 3 shows the results of the tests of between-subjects effects. A significant main effect was found to exist among the noticing scores of the three groups with an η^2 of 0.31, which is a rather large effect size according to Leech et al. (2005).

Post hoc analyses of the mean scores using Tukey test revealed a significant difference between EG2 and EG1 ($p < 0.05$) as well as EG2 and CG ($p < 0.05$). However, no significant difference was found between EG1 and CG ($p = 0.49$). Putting all these bits and pieces of information together, one could conclude that EG2 best benefited the treatment. In other words, the treatment yielded positive effects for EG2, but not so for EG1 and CG.

4.2 Results of the MANOVAs

Learning was measured in terms of the participants' scores on both the pretest and the posttest. Both of these tests included a recognition and a production section. First, the results of the comparisons between the pretest and posttest scores on the recognition items will be reported. Following that, the scores having to do with the production section will be dealt with.

4.2.1 Results of the MANOVA (Recognition)

In order to compare the three groups in terms of their pretest and posttest scores on the recognition items, a MANOVA (multi-variate analysis of variance) was run, the main assumptions of which were met, including *normality and outliers, linearity, and multicollinearity* (Pallant, 2001). Table 4 presents a summary of the descriptive statistics for the three groups on both the pretest and the posttest (the recognition section).

The results of the MANOVA turned out to be significant, Wilks' $\Lambda = 0.36$, $F(56, 112) = 18.10$, $p < 0.05$. However, follow-up univariate ANOVAs indicated that no significant difference was found to have existed among the groups on their pretest scores, $F(2,57) = 0.79$, $p = 0.45$. This non-significant result guaranteed the assumption of homogeneity of the groups at the outset of the study, before the treatment was given. Nevertheless, the treatment yielded significant results on the posttest, with $F(2,57) = 33.03$, $p < 0.05$. For further details of the statistical analyses, refer to Tables 5 and 6.

Table 5 presents the results of the multivariate tests for the recognition scores. A significant main effect was found to exist among the groups, as already reported, with a rather large effect size ($\eta^2 = 0.39$). Table 6 shows the univariate tests for the pretest and the posttest scores separately. As already reported, there was no significant difference among the groups as far as their pretest scores were concerned. However, after the treatment, a significant difference was found among the groups in terms of their posttest scores. The very large effect size also reassures us that the variability in the data was mainly due to the treatment given ($\eta^2 = 0.53$).

Post hoc analyses of the data revealed that there existed a statistically significant difference between EG1 and CG ($p < 0.05$) and also between EG2 and CG ($p < 0.05$), but not between EG1 and EG2 ($p = 0.97$). In this way, it could be claimed that the use of output tasks had positive effects for both EG1 and EG2, meaning that both of these groups outperformed the other one in the end. Therefore, output was found to have been effective in promoting the participants' learning of the target structure.

4.2.2 Results of the MANOVA (Production)

In order to compare the three groups in terms of their pretest and posttest scores on the production section, a second MANOVA was run, the main assumptions of which were also met. Table 7 presents a summary of the descriptive statistics for the three groups on both the pretest and the posttest (the production section).

The results of the MANOVA turned out to be significant, Wilks' $\Lambda = 0.38$, $F(56, 112) = 17.16$, $p < 0.05$. Nonetheless, follow-up univariate ANOVAs indicated that no significant difference was found to have existed among the groups on their pretest scores, $F(2,57) = 0.007$, $p = 0.99$. This non-significant result guaranteed the assumption of homogeneity of the groups at the outset of the study, before the treatment was given. In contrast, the treatment yielded significant results on the posttest, with $F(2,57) = 45.64$, $p < 0.05$. For further details of the statistical analyses, look at Tables 8 and 9.

Table 8 presents the results of the multivariate tests for the production scores. A significant main effect was found to exist among the groups, as already reported, with a rather large effect size ($\eta^2 = 0.38$). Table 9 shows the univariate tests for the pretest and the posttest scores separately. As already reported, there was no significant difference among the groups as far as their pretest scores were concerned. However, after the treatment, a significant difference was found among the groups in terms of their posttest scores. The very large effect size also reassures us that the variability in the data was mainly due to the treatment given ($\eta^2 = 0.61$).

Post hoc analyses of the data revealed that there existed a statistically significant difference between EG1 and CG ($p < 0.05$) and also between EG2 and CG ($p < 0.05$), but not between EG1 and EG2 ($p = 0.96$). This means that the use of output of tasks had positive effects for both EG1 and EG2. In other words, both of these groups outperformed the other one in the end. Therefore, output was once more found to have been effective in promoting the participants' learning of the target structure.

5. Discussion

5.1 Investigating the First Research Question

The first research question investigated the effect of using output tasks on learners' noticing of the past tense (preterit, progressive, perfect). In other words, the researcher intended to find out whether output tasks would actually be of any advantage or disadvantage to learners in promoting their noticing of the target structure. In order to do so, the statistical analyses explained above were carried out, and it was found that output tasks indeed made a positive difference in the participants' noticing.

However, one should be cautious in interpreting the results. This is because not both types of output tasks used in this study yielded positive results. Only reconstruction tasks were found to have significantly promoted noticing among the participants. In other words, this study has shown that task type is one of the parameters that can make or break successful noticing.

Placing the findings of the current study alongside those of previous studies in the literature, it can be said that this study (partially) corroborates the findings of Swain and Lapkin (1995), Izumi (2000), Ahmadi (2003), Hanaoka (2007), Song and Suh (2008), Kong (2010), and Ghari and Moinzadeh (2011).

One of the reasons why EG2 participants were more successful in promoting their noticing of the past tense probably lies in the nature of the task they were required to perform. It seems that reconstruction tasks make learners receive input with more focused attention to its formal features. The reconstruction task used in this study asked the learners to listen to an input text twice. During the first round of listening, they had to listen to the text in order to fully grasp it. Next, as they were listening to the same text for a second time, they were asked to take notes of any word or phrase that would help them later in reconstructing the text. It could be that the very nature of this output task may have prompted them to attend more to the grammatical features of input, including past-tense verbs, as they had to use them subsequently in their own writings without being given any verbal or visual clues.

In contrast, picture-cued writing tasks provide learners with both pictures and cues, based on which they can produce output. The fact that they have visual and verbal cues, by referring to which they can write their own versions of the text, probably causes them to attend lesser to the grammatical features of the input. However, such a difference between task types was neither observed in Song and Suh (2008), nor in Ghari and Moinzadeh (2011).

5.2 Investigating the Second Research Question

This research question explored the effect of using output tasks on the participants' learning of the past tense. That is to say, the researcher intended to find out whether output tasks would actually be of any advantage or disadvantage to learners in promoting their accuracy of using the target structure. To fulfill this aim, the statistical analyses explained above were carried out, and it was found that output tasks indeed had a positive influence on the participants' learning of the target form.

Unlike noticing, learning was equally improved among the two experimental groups, each of which was exposed to a different output task. It can be concluded that the use of output tasks, no matter which of the two types experimented with here, is effective in raising learners' grammatical accuracy as far as learning the past tense is concerned.

This finding is in line with those of previous studies in the literature which have come up with positive effects for the use of output tasks in teaching and learning grammar. This study (partially) corroborates the findings of Izumi et al. (1999), Izumi (2000), Mennim (2008), Song and Suh (2008), Kong (2010), and Abadikhah and Zarrabi (2011), and Ghari and Moinzadeh (2011).

It seems that in order to master a grammatical feature, merely being exposed to input containing the target form is not sufficient. One also needs to be given the opportunity to produce that form. In this way, he/she gets to know about the gaps in his/her interlanguage and will thus attend to subsequent input with more focused attention.

Also, continuous exposure to input along with requirements of output production could have caused the two experimental groups to outperform the comparison group in terms of their gained partial mastery over the use of the past tense. That is, according to the findings of this study, receiving output and producing input over a relatively extended period of time could hold the key to successful learning of grammatical features (Izumi & Izumi, 2004). Yet it is worth noting that the effect of length of time (over which input is received and output is produced) has to be separately investigated in future studies through comparing two or more groups with the same output tasks performed over different lengths of time.

Altogether, this study found evidence in support of Swain's claims about the functions of output, such as the noticing/triggering function. It also resulted in improvements in the learners' interlanguage.

6. Conclusions, Implications, and Suggestion for Further Research

The statistical analyses conducted revealed that output tasks of the reconstruction type were effective in promoting learners' noticing of the past tense. The same significant results were not found for picture-cued writing tasks, though.

In terms of learning, however, both output task types were found to have been effective in promoting learners' use of the target form, either recognizing or producing it. Altogether, output was shown to be of great benefits to learners in teaching grammatical features, particularly the past tense.

As already explained in the literature on noticing and language learning, successful learners are the ones who are able to notice the gaps in their interlanguage and who can receive subsequent input with more focused attention to the features which are problematic to them. Therefore, teachers are supposed to help their learners notice their language-related shortcomings if they would like to facilitate the learning process and foster their students' accuracy.

Keeping this in mind alongside the new findings of this study, it seems that teachers can play a role in helping their students overcome their problems through providing them with opportunities to make use of their linguistic resources. As a result, it is recommended that teachers employ output tasks as follow-up activities after teaching grammatical points which are specifically difficult for learners to use fluently and accurately. Subsequent exposure to target-like input is also necessary in helping learners compare their own interlanguage with that of native speakers. Such comparisons result in promoted noticing of linguistic forms.

The findings of the current study also have implications for material developers working in the area of grammar. They had better keep reminding themselves that what learners and teachers need is not merely more details of the grammatical features to be taught. Instead, after presenting each new point, some sort of follow-up activity has to be done, one which provides learners with chances to produce oral and written output and to put to use what they have just learned. This not only helps learners better master the form they are being taught, but also it benefits them in consolidating their knowledge of the previously-learned forms. In the same manner, teachers would not have to go through all those painstaking stages of correcting their learners' errors each time they occur. Rather, they could present their students with target-like well-formed models so that the students themselves would come up with their own defects in using the language they are learning. This way, learners are also encouraged in becoming increasingly autonomous in that they would not have to wait for their teachers to tell them about their errors, but they would notice them by themselves.

Future research studies can assess the effects of output treatment in relation with length of time, gender, age, target structures, and mode of production (oral vs. written output).

The researchers' other suggestion for future studies is relevant to those working in the area of validating measures. The literature on measuring noticing has documented several online and offline measures such as underlining (used in this study), think-aloud protocols, retrospective questionnaires, and interviews. However, it is not clear whether any of these measures or other similar ones actually measure learners' noticing of a given form. It would be beneficial to researchers if one could find out which method actually tests the construct under focus, i.e. noticing.

References

- Abadikhah, S., & Zarrabi, F. (2011). The effects of output tasks on the acquisition of English verbal morphemes. *Theory and Practice in Language Studies*, 1(11), 1549-1560. <http://dx.doi.org/10.4304/tpls.1.11.1549-1560>
- Bloor, M., Bloor, T., Forrest, R., Laird, E., & Relton, H. (1970). *Objective tests in English as a foreign language*. London: Macmillan.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209-224.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83-107.
- Ghari, A., & Moinszadeh, A. (2011). The effects of output task types on noticing and learning of English past modals: A case of intermediate Persian adult learners of English. *Journal of Language Teaching and Research*, 2(5), 1180-1190. <http://dx.doi.org/10.4304/jltr.2.5.1180-1191>
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459-479.
- Hill, L. A. (1980). *Anecdotes in American English*. Oxford: Oxford University Press.
- Izumi, M., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21(3), 421-452.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL

- relativization. *Studies in Second Language Acquisition*, 24(4), 541-577.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-278.
- Izumi, Y., & Izumi, S. (2004). Investigating effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review*, 60(5), 587-609.
- Kong, E. Y. (2010). Effects of output and note-taking on noticing and interlanguage development. *Working Papers in TESOL & Applied Linguistics*, 10(2), 19-36.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (1985). *The input hypothesis: Issues and Implications*. London: Longman.
- Krashen, S. (2009). The Comprehension hypothesis extended. In T. Piske, & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 81-94). Bristol: Multilingual Matters.
- Krashen, S., & Terrell, T. (1983). *The natural approach*. Oxford: Pergamon.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Interpretation and use*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Lightbown, P. M., & Spada, N. (2006). *How languages are learned*. Oxford: Oxford University Press.
- McLaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.
- Mennim, P. (2007). Long-term effects of noticing on oral output. *Language Teaching Research*, 11(3), 265-280. <http://dx.doi.org/10.1177/1362168807077551>
- Morgan, G. A., Leech, N. L., Cloeckner, G. W., & Barrett, K. C. (2004). *SPSS for introductory statistics: Interpretation and use*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Pallant, J. (2001). *SPSS survival manual*. Philadelphia: Open University Press.
- Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137-174). Rowley, MA: Newbury House.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010*, Singapore, December 2-4 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Song, M., & Suh, B. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System*, 36(2), 295-312. <http://dx.doi.org/10.1016/j.system.2007.09.006>
- Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer (Eds.), *Principles and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and Research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471-484). Mahwah, NJ: Lawrence Erlbaum.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391. <http://dx.doi.org/10.1093/applin/16.3.371>

Table 1. Noticing of the Target Structure (Based on Underlines)

	Group	Mean	Std. Deviation	N
Time 2	EG1	25.38	7.17	18
	EG2	23.78	7.69	19
	CG	22.60	8.31	23
	Total	23.81	7.74	60
Time 3	EG1	26.61	10.63	18
	EG2	29.68	11.68	19
	CG	24.69	10.53	23
	Total	26.85	10.95	60
Time 4	EG1	25.44	10.37	18
	EG2	34.10	13.90	19
	CG	24.52	10.50	23
	Total	27.83	12.23	60
Time 5	EG1	27.77	9.31	18
	EG2	42.26	9.40	19
	CG	25.91	10.67	23
	Total	31.65	12.17	60

Table 2. Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Sessions	Greenhouse-Geisser	1968.53	2.21	890.50	6.47	.001*	.10
Sessions * Group	Greenhouse-Geisser	1770.96	4.42	400.56	2.91	.020*	.09
Error(Sessions)	Greenhouse-Geisser	17338.24	126.00	137.60			

a. Computed using alpha = .05

b. * = significant

Table 3. Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Group	2836.89	2	1418.44	12.92	.000	.31
Error	6256.51	57	109.76			

Table 4. Descriptive Statistics: Pretest/Posttest (Recognition)

	Group	Mean	Std. Deviation	N
Pretest	EG1	12.16	1.33	18
	EG2	12.05	1.17	19
	CG	11.69	1.25	23
	Total	11.95	1.25	60
Posttest	EG1	16.38	1.14	18
	EG2	16.47	1.30	19
	CG	13.56	1.47	23
	Total	15.33	1.91	60

Table 5. Multivariate Tests (Recognition)

Effect		Value	F	Error df	Sig.	Partial Eta Squared
Intercept	Wilks' Lambda	.00	4.118E3	56.00	.000	.99
Group	Wilks' Lambda	.36	18.101	112.00	.000*	.39

a. Computed using alpha = .05

b. * = significant

Table 6. Tests of Between-Subjects Effects (Recognition)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
Group	Pretest	2	1.26	.79	.455*	.02
	Posttest	2	58.33	33.03	.000**	.53
Error	Pretest	57	1.58			
	Posttest	57	1.76			

a. Computed using alpha = .05

b. * = significant

c. ** = insignificant

Table 7. Descriptive Statistics: Pretest/Posttest Production

	Group	Mean	Std. Deviation	N
Pretest	EG1	10.70	1.56	18
	EG2	10.64	1.56	19
	CG	10.68	1.45	23
	Total	10.68	1.49	60
Posttest	EG1	15.45	1.57	18
	EG2	15.32	1.38	19
	CG	11.47	1.63	23
	Total	13.88	2.44	60

Table 8. Multivariate Tests (Production)

Effect	Value	F	Error df	Sig.	Partial Eta Squared
Wilks' Lambda	.00	3.897E3	56.00	.000	.99
Group	Wilks' Lambda	.38	17.16	.000*	.38

a. Computed using alpha = .05

b. * = significant

Table 9. Tests of Between-Subjects Effects (Production)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
Group	Pretest	2	.01	.007	.99	.00
	Posttest	2	108.41	45.64	.00	.61
Error	Pretest	57	2.32			
	Posttest	57	2.37			

a. Computed using alpha = .05

b. * = significant

c. ** = insignificant

Appendix

Sample Picture-cued Writing Task from the First Session

Look at the pictures and try to rewrite the story using the phrases given.



get out of a train



give coins to a beggar



meet another beggar



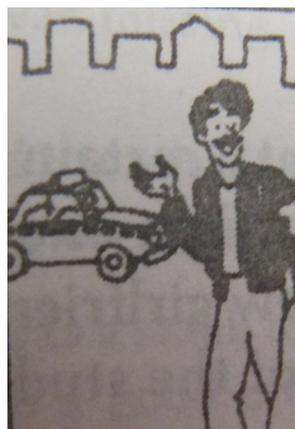
take the beggar to a restaurant



bring the bill / not pay it



pay the bill



take the beggar to a taxi



refuse to get in

An artist _____

Sample Reconstruction Task from the First Session

Listen to the text for a second time and write down any words or phrases that will help you in rewriting the text.

DO NOT WRITE COMPLETE SENTENCES, but just words or groups of words.

Now try to rewrite the story in your own words using your notes.

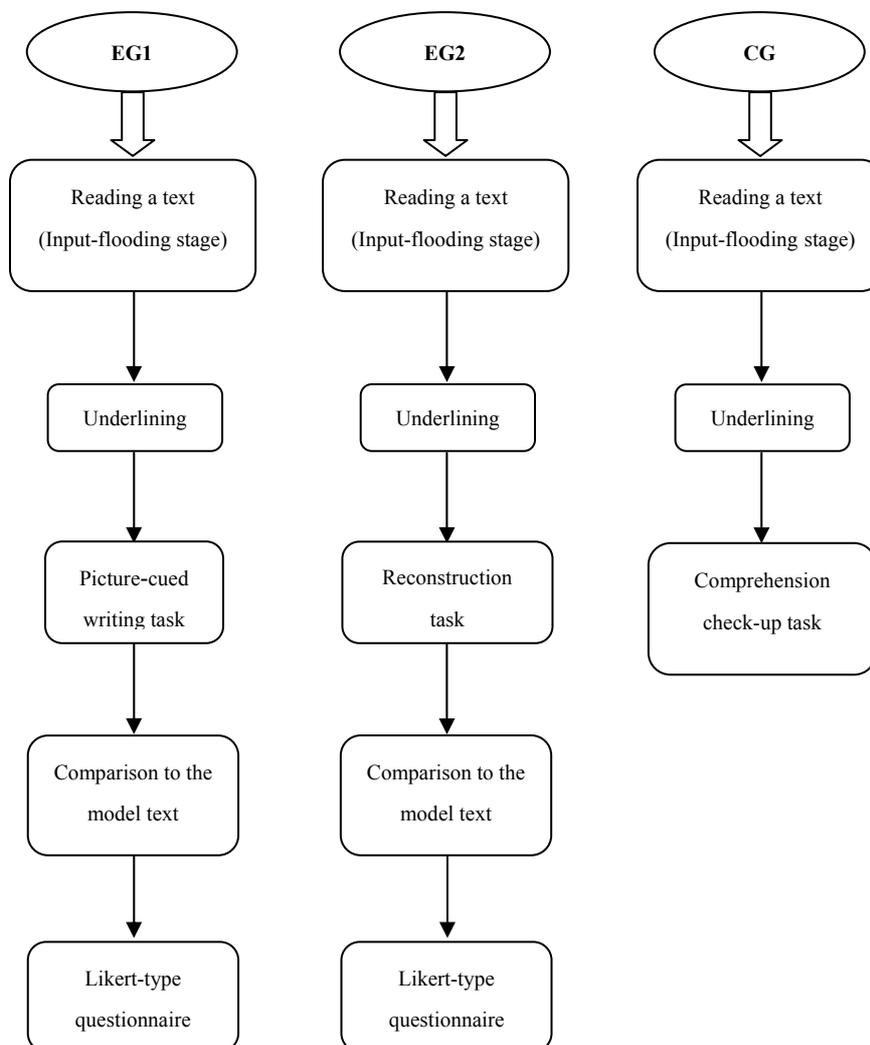


Figure 1. Overview of a single treatment session for each of the three groups