Differential Potential of SLA Output Tasks versus Input-based Teaching of English Grammar: A Comparative Study

Mehraban Hamavandy
Department of English, Tarbiat Modares University, Tehran, Iran

Mohammad Golshan
Department of English, Maybod Branch, Islamic Azad University, Maybod, Iran

Abstract—The present study is an attempt to investigate a) the effect of output requirement on the acquisition of grammar and b) the differential functions of two output tasks (picture-cued production and editing reconstruction tasks) on the interlanguage improvement of participants. To this end, twenty four Iranian elementary EFL learners were assigned to two experimental groups (EG1 & EG2) and a control group (CG). A pre-test was administered to the three groups to measure their knowledge of the target grammar structure (present perfect tense) prior to any treatment. The participants in EG1 were required to produce their language based on a set of pictures presented, while the learners in EG2 were asked to edit a given text based on its grammatical appropriacy. Those in CG were exposed to the target grammar item merely by providing input. The effectiveness of the treatment tasks was measured using a post-test. The results revealed that those participants in the CG who were provided with input outperformed their peers in the EGs. Also there was no statistically significant difference found between the two output tasks.

Index Terms—output hypothesis, output task, input-based instruction, teaching grammar

I. INTRODUCTION

Recent findings in second language acquisition (SLA) studies point to the fact and make it clear that ‘noticing’ is a leading factor in SLA development (Ellis, 1994; Long, 1991; McLaughlin, 1990; Robinson, 1995; Schmidt, 2001; Smith, 1993; Tomlin & Villa, 1994). For a long time, the process of language learning was deemed to be merely input-based, thus calling for mechanisms to enhance the noticing of the input to turn it into intake, which would, in turn, improve the interlanguage of the learners (Gass, 1997; Johnson, 1996; Skehan, 1998; Van Patten, 1993). Later on, it was hypothesized that in addition to input (intake), it is output that can function as a key factor in acquisition of language, establishing it as a legitimate approach in teaching different language skills, including grammar. (Schmidt, 1992; Swain 1985, 1995, 2000)

The bulk of studies conducted thus far have primarily investigated different methods of providing output production to enhance the noticing of learners (Doughty, 2001; Doughty & Williams, 1998; Izumi, 2002; Izumi et al., 1999; Izumi & Bigelow, 2000; Long, 1991). The underlying assumption of these studies has been that drawing learners’ attention to the target linguistic forms, which might occur incidentally during communication, facilitates acquisition of both form and meaning. Yet, a significant issue which has not received due attention in many of the studies conducted on the role of output as a promoting factor of learners’ interlanguage (IL) is the comparison of different production tasks which might yield different results in different domains of language, including grammar, vocabulary, and pragmatic acquisition. This study is an attempt to compare two output tasks, namely editing reconstruction and picture-cued production tasks, with regard to grammar improvement for learners of English as a foreign language in the context of Iran.

II. BACKGROUND

In SLA literature it was a universally accepted and long established belief that comprehensible input plays the dominant role in improving the interlanguage of the foreign language learners (Krashen, 1985; Loschky, 1994). In other words, the proponents of ‘comprehensible input theory’ maintained that language learners can improve if provided with ample input which is both comprehensible and slightly beyond the current linguistic level of the learners (i+1).

In 1985 Merrill Swain who is considered, by many, as the initiator of what is now known as the ‘Output Hypothesis’ argued that, contrary to what was the prevalent belief then, only comprehensible input could not benefit the interlanguage of nonnative learners of a language. In other words, she maintained that providing opportunities for the learners to produce the language was a crucial factor to improve the language proficiency of the learners. Her main sources of evidence were findings from immersion programs in Canada, where she observed that unlike plenty of
comprehensible input provided many of the learners lagged far behind the expectations in terms of grammatical and sociolinguistic competence. In a number of studies (Allen, Swain, Harley and Cummins, 1990; Harley & Swain, 1978) it was shown that although learners in immersion programs were fairly proficient in their discourse skills and were confident enough to use L2, they lacked full sociolinguistic competence and were unable to master more marked grammatical distinctions. This could not be justified by lack of sufficient comprehensible input of which immersion programs were rich and ample practice had been done previously. Swain proposed her comprehensible output hypothesis to signify the role of the learners’ output in SLA studies. She defined comprehensible output as “the output that extends the linguistic repertoire of the learner as he or she attempts to create precisely and appropriately the meaning desired” (Swain, 1985, p. 252). She contended that production entailed syntactic (bottom-up) processing, while comprehension relied mainly on semantic (top-down) processing. Thus, concluding that while comprehension of a message might occur with little syntactic analysis of the input, production requires attention to the means of production, especially if learners are to produce socially appropriate messages.

Swain proposed three functions for comprehensible output of the learners: 1. noticing (consciousness-raising) function, 2. hypothesis-testing function and 3. metalinguistic function. In her later modifications of the theory, Swain added a fourth component of ‘fluency function’ which deals with the ability to develop quick access to the language repertoire for a fluent productive performance.

Evidence from studies conducted to investigate the effectiveness of providing opportunities for production of L2 learners is mixed, and far from being taken as conclusive. For instance, Ellis (2008) contends that there is no evidence which can unambiguously show the relationship between learners’ productive participation in the classroom and its impact on the rate of language development. He further elaborates on the results of many studies that have been conducted on the relationship between learner production and language proficiency and concludes that yet the results are mixed and inconclusive (p.807).

Shehadeh (2002), in a similar vein, believes that:

After well over a decade of research into Swain’s (1985) comprehensible output hypothesis, few definitive conclusions can be made, because the question of whether and how learners’ output, or output modification, helps with L2 learning is still largely unanswered (p. 601).

Izumi and his colleagues researched the noticing function of comprehensible output in a number of their studies. The predominant thrust of their studies was to investigate the role of output as a prompt for noticing followed by learning certain targeted grammatical features. (Izumi, 2002; Izumi & Bigelow, 2000; Izumi & Izumi, 2004; Izumi, Bigelow, Fujiwara & Fearnow, 1999). For instance, Izumi et al. (1999) investigated the noticing function of output and its probable result on the performance of the participants. To this end, they addressed two research questions: (a) does output promote noticing of linguistic form? and (b) does output result in improved performance on the target form? The participants were asked in phase one of the treatment to reconstruct a short passage, followed by a second reconstruction after being exposed to it, and in phase two to write about a given topic twice, just prior to their second writing they were presented with a model written by a native speaker. In order to check the noticing function of the output, the participants were asked to underline parts of the sentences they thought might be “particularly necessary” for subsequent (re)production. The control group only received input material. The results indicated that output production could not provide any significant effects on noticing of the form. On the other hand, the experimental group significantly outperformed the control group on the production task.

In another study Izumi and Izumi (2004) studied the role of providing language learners with the opportunity to have oral output and its effect on the ultimate gains of the participants in comparison with non-output-provided learners. They randomly assigned the twenty-four participants into the three groups of: output (engaging in a picture description task involving input comprehension and output production), non-output (engaging picture description task requiring input comprehension only), and a control group. The findings indicated that, to the surprise of the researchers, the non-output group revealed to have more learning gains than the output group. The authors attribute the results to the different cognitive processes that the participants probably underwent in different groups. In the output group the participants only engaged in oral production which seems less cognitively demanding in terms of syntactic processing than picture sequencing task conducted by the non-output group. This might have facilitated the ability of form-meaning mappings of the subjects.

In a similar vein, Horibe (2002) in his study, on the role of providing output for learning several syntactic structures (target forms), compared two instructional treatment conditions (input only and input-output). The subjects were 31 college students in 3 intact classes of input only (input group), input and output (output group), and a placebo group receiving no instruction (control group). The subjects’ thought processes in the spoken output were elicited and recorded via think-aloud protocol interviews. The results revealed an insignificant difference between the two groups of input vs. output in terms of learning the target structures.

Shehadeh (2002) explored learners’ “hypothesis testing episodes” during a picture description task. Each participant was aided by a native speaker partner in the study. Learners were required to monitor and modify their initial output, by confirming the structural correctness with the native speaker. Shehadeh concludes that learners confirmed their self-initiated hypotheses as correct when did not received any feedback from the native speakers.
In another study, Ellis and He (1999) conducted a research on different effects of giving input and providing output on language proficiency of English learners. Their experiment was based on a listen-and-do task regarding four different conditions of: a) unmodified input b) pre-modified input c) interactionally modified input and d) modified output. The findings were in line with Swain’s hypothesis. They revealed that those participants who were required to produce the comprehensible output (fourth group) outperformed those who received input only (in its different forms). In this vein of research, De La Fuente (2002) observed that learners of Spanish vocabulary who were involved in negotiated interactions that entailed pushed output were, to a far extent, superior in productive acquisition of new words than those who received either pre-modified input or negotiated interaction without pushed output.

In sum, due to the inconclusive evidence of the studies conducted by many researchers on the role of output in SLA literature, there seems to be justification for another study on the issue. As Ellis (2008) puts it:

What is not yet clear, however, is whether output assists learners to acquire new linguistic forms or only to automatize use of partially acquired forms. Further work is needed to establish whether (and under what conditions) the modified output constitutes acquisition.

In this study the primary aim is to investigate the differential roles of output tasks on the acquisition of a certain grammatical item (present perfect tense) by Iranian learners of English as a foreign language. Specifically, this study is guided by the following two questions and seeks answers for them:

1. Does output requirement result in a significant difference on learning grammar over merely providing input for the learners?
2. Is there a significant difference between the two output tasks of editing and picture-cued with regard to acquisition of grammar?

III. METHODOLOGY

A. Participants

The participants of this study were 45 Iranian Elementary learners of English in one of the language institutes in Iran, namely Iran Language Institute. They ranged in age from 16 to 28, and were, minimally, passing the courses at the institute for 1.5 years. They were assigned to either groups of experimental group 1 (EG1), experimental group 2 (EG2), and control group (CG) on a random sampling basis, and were offered an extra-curricular training for participation in the study.

B. Target Form on Focus

The target structure of this study was the present perfect tense (e.g. I have studied my lessons for two hours since ten o’clock). This tense was deemed to be teachable (Pienemann, 1989), since the learners were already familiar with other tenses of present, past and present continuous. The participants’ attempts to use this tense in the pre-test of the study revealed that they were not familiar with it, and could not use it accurately.

C. Instrumentation

In order to observe the gains of the participants a test which aimed to tap the present perfect tense knowledge of the learners was developed. The test consisted of nine items whose main verbs were deleted (acting as the blanks of the test totaling 12 in number). The participants had to produce the correct form of the verb with regard to the tense of the sentence. The interval between the pre-test and the post-test was a six-week period intervened by the treatment of the study Also, in order to ensure the internal consistency of the test it was measured using Cronbach’s α which turned out to be 0.83 for the pre-test and 0.76 for the post test.

D. Data Collection Procedure

This experimental study was carried out over a period of around eight weeks. As stated earlier there were 45 participants in this study, who were assigned to the three groups of EG1, EG2, and CG. One week prior to the initiation of the treatment the pre-test was administered to the participants. Then the two groups of EG1 and EG2 underwent the treatment of the study which continued for eight consecutive sessions and entailed doing picture-cued production tasks for EG1 and editing reconstruction tasks of English grammar tenses for EG2 (see Appendices A and B for samples of the treatment). Using tasks has revealed to be a productive means of investigating classroom research. ‘Task’, as a general overarching term, has been figured as a good device for delivering instructional treatment in experimental studies, and for measuring the outcomes of this treatment. Therefore, as Pica (1997) contends, tasks can serve as a bridge between pedagogy and research in SLA studies.

Both groups of EG1 and EG2 were briefly introduced into the functions of present perfect tense prior to receiving the treatment. The CG, as the placebo group, only received input-based lessons of present perfect tense which were according to the teaching principles of the institute course book.

As for the participants in EG1, they were required to produce a short text (at a minimum of three or four sentences) describing the pictures shown to them. It was ensured that the participants could use a variety of tenses in their writings, including the present perfect tense.
The texts provided for the participants, to be edited, in EG2 included three to five sentences with underlined verbs. The participants were required to check the tense of the verbs in the sentences and correct them if used inappropriately. The sentences were of a variety of tenses including present, past and present perfect tense. Storch (1997) concluded in his study that an editing task succeeds in drawing learners’ attention to a range of grammatical and lexical choices targeted for instruction. The tasks were given to the students in the last half hour of the class (a one hour and forty-five minute class). The students completed the tasks individually and could leave the class upon completion of the task. The whole procedure started in session four of the term and ended in session twelve.

IV. DATA ANALYSIS

In scoring the pre- and post-tests of the study each correct response received one point and all the scores were added up to a total of sum. There was no penalty assigned for the wrong responses. In order to compare the control and the experimental groups’ scores on the pre-test, and post-test, two independent sample t tests were run, using SPSS 21, setting the level of significance at 0.05.

V. RESULTS

Analysis of the Pre-test of the Study

For such a study to sound meaningful and also to check for the comparability of the knowledge of the participants regarding the target structure (present perfect tense) a pre-test was administered. The descriptive results of the pre-test are shown in the Table 4.1 below.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG1</td>
<td>12</td>
<td>3.00</td>
<td>12.00</td>
<td>7.4166</td>
<td>3.5791</td>
</tr>
<tr>
<td>EG2</td>
<td>12</td>
<td>1.00</td>
<td>12.00</td>
<td>6.5454</td>
<td>3.5032</td>
</tr>
<tr>
<td>CG</td>
<td>21</td>
<td>2.00</td>
<td>12.00</td>
<td>6.4583</td>
<td>1.7687</td>
</tr>
</tbody>
</table>

As can be observed from the table above the mean scores of the three groups are very close to each other (7.41≈ 6.54≈ 6.45), thus convincing that the prior knowledge of the structure in question was almost statistically equal for the three groups of the study.

Research Question 1

The first research question aimed to investigate whether there is any significant difference between the final target structure attainments of the participants in the output-required versus input-based groups. In other words, the result of the post-test of the CG was to be compared with that of the EGs. In order to come up with the result, an independent sample t-test was run to measure the significance of differences between the means of the two groups. The results are displayed in Table 4.2 and Table 4.3 below.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG</td>
<td>24</td>
<td>5.91</td>
<td>2.21</td>
<td>.46</td>
</tr>
<tr>
<td>CG</td>
<td>21</td>
<td>7.31</td>
<td>1.98</td>
<td>.42</td>
</tr>
</tbody>
</table>

Based on the information given in the Table above, it can be seen that the mean of the results of the post-test for the CG (7.31) is higher than that of the EG (5.91).

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>Df</th>
<th>Sig. (2 tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.36</td>
<td>44</td>
<td>.023</td>
<td>-1.42</td>
<td>0.60</td>
<td>[-2.63, -0.20]</td>
</tr>
</tbody>
</table>

Based on the information given in Table 4.3 it can be deduced that there is a significant difference in the scores obtained from the CG and EG because the probability value is smaller than the specified critical value (0.023<0.05) and the t statistic is at a high value (2.36).

Thus it can be claimed that the input-based only teaching of grammar for the participants of CG helped them significantly outperform those learners who were attending output-required classes of EG.

Research Question 2

The second research question asked whether there is any significant difference in the final performance of the learners who attended the experimental groups and were required to have production during the treatment sessions, namely EG1 who did picture-cued tasks, and EG2 who underwent editing tasks. In order to answer this question an
independent-sample t test was run. Tables 4.4 and 4.5, below, provide the descriptive statistics along with the results of the independent-sample t test.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>S.E. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG1</td>
<td>12</td>
<td>5.91</td>
<td>2.57</td>
<td>0.74</td>
</tr>
<tr>
<td>EG2</td>
<td>12</td>
<td>5.92</td>
<td>2.01</td>
<td>0.53</td>
</tr>
</tbody>
</table>

As can be seen from the results of Table 4.4, the mean scores of the two groups (EG1 and EG2) revealed to be surprisingly near to each other (EG1 = 5.91 ~ EG2 = 5.92). Also, on a closer inspection of the t-test on Table 4.5 the probability value (0.99) was observed to be bigger that the critical value (0.05) signifying the similarity of performance of the two groups. In sum, it can be maintained that the two different output tasks administered on the participants could not exert any differential influence on the ultimate acquisition of the target structure.

Furthermore, in order to see where the difference(s) among the means of the three groups may lie, a post-hoc Scheffe test was run on the post test results. (Table 4.6). The results of Scheffé’s test indicated that there were not significant differences between the means of the three groups (CG, EG1 and EG2). This is, vividly, indicative of non-difference in performance of the subjects in either input-received or output-required groups.

### VI. DISCUSSION

This study was an attempt to answer the two questions which have been regularly addressed in SLA studies, namely comparison of output- versus input-based learning of a language, and second, comparison of different production tasks for language learning. Results from the first question of the study revealed that input-based learning did influence the learning of the target grammatical structure in a more significant way than the output-based counterpart. While this finding might seem in direct contrast with the tenets of output hypothesis and many of the studies conducted on the issue (Izumi, 2002; Izumi et al., 1999; Song & Suh, 2008; Swain 1993, 1995; Swain & Lapkin, 1995), yet it is in line with other studies (Horibe, 2002; Izumi & Bigelow, 2000; Izumi & Izumi, 2004; Shintani, 2011) which have specifically compared the input vs. output-based treatments in language acquisition and came up with the finding that output-based teaching could not be deemed any superior to the input-based equivalent. These studies point to the non-significance of differences between learning gains of participants involved in output tasks in comparison with those engaged in non-output tasks. Though the account advanced here implies an ineffective role of providing output opportunities through different production tasks, this does not negate the probable suitability of such practices for different learners with different linguistic, social and psychological backgrounds.

As of the second research question there was no significant difference found between the two output tasks of editing reconstruction and picture cued. This is not consistent with the general trend observed in earlier studies. For instance Nassaji and Tian (2010) as part of their study compared two output tasks, and came up with the finding that reconstruction editing task was more effective in terms of promoting negotiation and learning than reconstruction cloze task, Storch (2001) also in his study which compared the performance of tertiary ESL learners of intermediate second language proficiency on three different grammar-focused classroom tasks found that although all three tasks succeeded in drawing the learners’ attention to a range of grammatical items, the text reconstruction task was the most successful in doing so. The participants of the study also reached correct grammatical decisions in a majority of instances. Thus it could be contended that different output tasks may yield different influences on language learning assuming the differences between language learners and forms.
In sum, based on the inconclusive, and at times controversial, results obtained thus far further research on the issue, also taking into account learner differences seems necessary.

VII. CONCLUSION

This study revealed as its findings that first, input-based teaching for grammar does a better job than providing output tasks for the learners and, second there is no difference between the two output tasks of picture-cued production and editing reconstruction. Thus, adding to the confusion over the preferred way to enhance the learners’ noticing and ultimate attainment of the language. In terms of output tasks this piece of research suggests a further inquiry into the issue for different learning outcomes stimulated by various task types. To further probe the application of output in noticing enhancement and SLA, future studies are recommended to explore various grammatical forms under varying circumstances. This vein of research in SLA is also hoped to shed light on the conditions under which output production in line with input enhancement can facilitate and promote the learning of a second language.

APPENDIX A. SAMPLE PICTURE-CUED PRODUCTION TASK

Directions: Please look at the picture and produce a text of yours (consisting of at least one paragraph to describe the event of the picture). Use the verb given in your composition.

![Burn Image]

APPENDIX B. SAMPLE TEXT OF EDITING RECONSTRUCTION TASK

Direction: Put the verb in brackets into the correct tense, if necessary.
1. He **lived** (live) in London for two years and then **goes** (go) to Edinburgh.
2. When I left school I **have cut** (cut) my hair and **wore** (wear) it short ever since.
3. My brother **has written** (write) several plays. He **just finished** (just/finish) his second tragedy.
4. I **did not see** (not see) him for three years, I wonder where he is.
5. He **has not smoked** (not smoke) for two weeks. He is trying to give up.

REFERENCES


© 2015 ACADEMY PUBLICATION
include a number of refereed articles in applied linguistics journals.

- Mehraban Hamavandy

Mehraban Hamavandy is a PhD graduate of applied linguistics from Tarbiat Modares University, Tehran, Iran. His main areas of interest include foreign language teaching and assessment, program evaluation, and curriculum development. His publications include a number of refereed articles in applied linguistics journals.

© 2015 ACADEMY PUBLICATION
Mohammad Golshan is Assistant Professor in the Department of TEFL, Islamic Azad University, Maybod Branch, Iran. In addition to articles on SLA, corrective feedback, and learning strategies, he has authored a number of books on ELT.