Empirical Study on the Effects of Noticing in EFL*

Weiwei Pan

National Research Center for Foreign Language Education in BFSU, 100089, Beijing, China; Teachers' College of Beijing Union University, 100011, Beijing, China Email: panweiw@163.com

Abstract—Following an initial investigation (Izumi, Bigelow, Fujiwara, & Fearnow, 1999), this study aims to test the noticing function of output, produced by Swain (1993, 1995, 1998). Participants were divided into two experimental groups and one control group. The first experimental group was given opportunities for output, the second experimental group engaged in comprehension-based activities, and the control group performed nothing during this phase. The results of this study indicate no unique effects of output compared with answering the true-or-false comprehension questions. A closer examination of the data collected from an interview, which is aimed to find out the participants' attentional focus during the treatment, further suggests that retelling as a form of output does not always succeed in drawing the learners' attention to the target form, but to other aspects such as vocabulary; phrases in the input material.

Index Terms—output, noticing function, EFL

I. INTRODUCTION

Focus-on-form treatment, has received some theoretical support from SLA literatures. A general finding by the researchers (Robinson, 1995; Schmidt, 1990; and Truscott, 1998) indicates that attention is necessary for learning to take place. That is to say people learn about the things that they attend to and do not learn much about the things they do not attend to. .

Schmidt (1990, 1995) and Schmidt have proposed the Noticing Hypothesis. Schmidt and Frota (1986, p. 311) offer a "notice the gap principle" which is aimed to state " a second language learner will begin to acquire the target like form if and only if it is present in comprehended in input and `noticed' in the normal sense of the word, that is consciously". Schmidt claims that noticing is the necessary and sufficient condition for the conversion of input to intake for learning (as cited in Izumi & Bigelow, 2000). Gass (1988) also asserts that noticing is the first stage of language acquisition, and Sharwood Smith (1981) and Rutherford (1987) advocate that noticing a feature in the input is an essential first step in language processing.

Noticing, according to Schmidt, requires focal attention and awareness on the part of the learner, and subliminal learning cannot account for SLA processes (as cited in Izumi & Bigelow, 2000). Intake, according to Schmidt, is that part of the input that the learner notices. It means that input indicates perceived information whereas intake indicates the information which is noticed by the learner. Based on this distinction, he proposes that intake, or noticing linguistic forms, is critical to subsequent processing of the forms.

Noticing is of considerable theoretical importance because it accounts for which features in the input are attended to and so become intake (information stored in temporary memory which may or may not be subsequently accommodated in the interlanguage system). Schmidt and Frota (1986) suggest that for noticed input, to become intake, learners have to carry out a comparison of what they have observed in the input and what they themselves are typically producing on the basis of their current interlanguage system. They refer to this as "noticing the gap", and argue that this too is a conscious process.

The present study investigates the potentially facilitative effects of internal attention drawing device-output-on the acquisition of past hypothetical conditional in English by second language learners. Of the four functions of output specified in the output hypothesis, the present study focuses on the noticing function. The most important reason of our focus on the noticing function of output is that this function of output has significant theoretical and pedagogical implications. In theory, the noticing function of output is closely related to the issue of noticing in SLA; in pedagogy, although learner's output is a prevalent feature of many language teaching practices, yet, to be exact, whether and how it helps with language learning has often been assumed rather than vigorously tested and investigated. Therefore more empirical investigation in this area is urgently needed.

II. THE PRESENT STUDY

^{*} The paper serves as the research finding of the project "The Whole Brain Education and Evaluation". Issue No.of the research project is ZLH12001.

Research Hypothesis

This study is a partial replication of Izumi et al.'s (1999) study. However, his participants are two college-level academic writing classes with different first language backgrounds, whereas our participants are first-year college students in China. In Izumi et al's (1999) study, target form is not explicitly explained in any form in the input material, whereas in the present study, the target form imbedded in the input is given prominence by using measures such as bolding or underlining, and also explicitly explained in written form. Izumi is interested in whether output would promote noticing of linguistic form as well as whether output resulted in the improved performance on the target form. We, however, design two experimental groups differing with regard to treatments, and aim to decide which treatment does better in noticing the target form and consequently is more conducive to SLA.

Two research hypotheses in this paper are generated as follows:

Hypothesis 1: Output in the form of retelling will promote noticing of formal elements in the target language input.

Hypothesis 2: Experimental Group 1, which is required to retell a passage, will perform better in noticing the target grammatical form contained in the input than Experimental Group 2, which does not produce output but instead answers true-or-false comprehension questions.

III. METHODOLOGY

1. Target Form

The past hypothetical conditional in English is the target form for this study (e.g., You wouldn't have caught cold if you had put on more clothes.). Celce-Murcia and Larsen-Freeman (1983) note that conditional sentences in general, and hypothetical or counterfactual conditionals often cause problems for many ESL learners. Conditional sentences consist of two clauses: an if-clause and a main clause. The structure is quite complex because mastery of this structure requires a good grasp of the English tense system and the modal auxiliaries. This is the reason to choose this target form in the present study. Besides, all the participants have demonstrated in the pretest that they have gaps when they use the target form. So it suggests that this form is not completely new to them, but all of them still have problems using the form accurately.

Some examples of the multiple choices that are produced by the participants in the pretest are as follows:

(1) If you D- a peep hole in your door you would have seen who was standing

outside and kept the door shut.

A. have B. have been having C. had had D. have had

(2) I was too busy yesterday, otherwise I B to see you.

A. would come B. would have come C. come D. came

(3) If I B rubbed gloves on I would have been electrocuted.

A. had not had B. have not had

C had not been having D. have not been having

Such emergence of the knowledge of target form may indicate a certain readiness to learn the form, because many of the participants already attempt to use it in a meaningful way.

2. Participants

Participants for the study (N=117) were recruited from among students enrolled in the ESL programs at a university in China. The selection of the participants was determined on the basis of a test of the past hypothetical conditional in English. This test also served as the pretest for the participating subjects. 117 participants took the pretest, and 92 satisfied the teachability requirements. This requirement, framed as the Teachability Hypothesis by Pienemann (1989), holds that teachers must be familiar with the order and sequence of acquisition that learners in general manifest and with the developmental stage that individual learners have reached.

In this study, 90 participants agreed to participate in the study and completed the following treatments and the two posttests. Participants were assigned to two experimental groups (EG1, N1=30; EG2, N2=30) and a control group (CC; N3=30), using a random assignment procedure. Experimental groups completed all the treatments and posttests and the control group only received the treatment of input, reading the text without any task and the final posttests.

3. Testing instrument

Multiple-choice recognition test: In both the pretest and posttests, multiple-choice recognition test in the written form was used. The test consisted of 10 target items and 10 distractors in a multiple-choice format. Five of the target items began with an if-clause, and five with a main clause. A blank targeting a target form is left either in the if-clause or in the main clause.

Retrospective interview: In order to find out the participants' attentional focus during the treatment, an interview was administered immediately after the delayed posttest. 5 participants were selected randomly from each experimental group to receive the interview.

4. Procedure

Experimental Sequence

TABLE 1
Day1: Pretest (30 minutes)
Day3: Phase 1: Receiving explicit written form of input (10 minutes)
Phase 2:
EGI: Reading the text embedded with the target form, and retelling the text (30 minutes)
EG2: Reading the text embedded with the target form, and answering
true-or-false comprehension questions (30 minutes)
CG: Reading the text embedded with the target form, but performing nothing (30 minutes)
Immediate Posttest (30 minutes)
Day 16: Delayed Posttest (30 minutes)
Interview (60 minutes)

The experimental schedule of the .study is as follows. First, the researcher visited each class and administered the pretest. On the second day, those participants whose scores in the pretest satisfied the requirement were chosen for the study. During this day, EG1 were requested to come to the computer laboratory outside of class hours in order to familiarize themselves with the machine of the lab room, because EG1 were expected to do the retelling task in the following treatment, and they needed to know how to operate the recording system.

On the third day, two phase tasks were completed. EG1 were led to the lab room, and EG2 and CG were in another two separate rooms. Two researchers were allocated in each room to supervise the whole process of the experimental study.

In phase 1, all the participants in the three rooms were given a sheet of input material, containing the use of grammatical structure of past hypothetical conditional in English, with some examples below each annotation of grammatical rule. This input knowledge was not be explained by the researcher, but conveyed explicitly by the handout in which the target structure was enhanced by bolding or underlining. Participants could pay attention to whatever words, phrases, or grammar points they would like, and researchers did not give them any instructions or inform them of any of the following treatments. It took approximately 10 minutes for participants to study the input material, and then the input sheets were taken away by the researcher.

In the phase 2 task, all the participants were given a short passage in which approximately 40% of the sentences contained the past hypothetical conditional form.

For EGl, participants were informed in advance that they would have only a limited amount of time-about 15 minutes to read the text, and about 15 minutes to retell the passage. During the first 15 minutes, they could take notes or underline whatever they believed necessary, and then the passage was taken away by the researcher. After that, EGl were told to perform the retelling task and at the same time they needed to record what they said using the tape recorder in the lab room within 15 minutes.

For EG2, half an hour was given to complete to whole phase. All the participants were asked to read the same passage as EGI's, and they could underline or make any marks on the parts they felt important or necessary for answering the true-or-false comprehension questions. Then EG2 answered those questions written below the passage. When they did the true-or-false comprehension questions, they could refer to the reading material.

For CG, they had 30 minutes to read the same passage, and they could underline whatever part they believed important, but they had no other task. During the phase 2 of the study, participants of 2EGs and CG were required not to have any discussions about the reading material.

This study used retelling as the output treatment, similar to the text-reconstruction task used in the studies of Izumi and Bigelow (2000) and Izumi et al. (1999). There is one significant reason for using the retelling task, that is, retelling task by its very nature has the potential for promoting comparisons between the IL output and the TL input. It is essentially a meaning-based pedagogical activity that allows the participants to devote their attention to the target form. So it is a very important task in light of the notion of focus on form, in which integration of form and meaning is emphasized. In particular, the advantage of the retelling task lies in its control over the content and form that the participants produce. Maximizing the equivalence between participants' output and the target input should promote direct comparisons between their IL output and the TL input forms.

After the treatment period, the researchers administered the immediate posttest. The mean length of the immediate of posttest was about half an hour. After the immediate posttest was completed, the delayed posttest was performed 2 weeks later: It still took approximately 30 minutes, and then we randomly selected and interviewed 5 participants from each of the two EGs to obtain retrospective data '(see Appendix 6) on the cognitive processes engaged in different treatments. The interview took about an hour.

In order to control the outside exposure to the target form, the teachers of the participating classes were requested not to teach past hypothetical conditional during the experimental period, in particular during the period between the immediate posttest and the delayed posttest. All teachers accepted and followed this request. Participants were asked not to discuss the treatment activities with other people.

IV. DATA ANALYSIS

The answer to each target item was scored as either correct or incorrect, and the 10 distractors were eliminated when we counted the scores. After scoring the pretest and immediate as well as delayed posttests, gain scores were calculated

for each participant from the two posttests in order to examine how much improvement had been made on the target form. In the results section that follows, the test scores and the statistical results are reported.

Statistical analysis was carried out with the aid of Statistical Product and Service Solutions (SPSS 11.0). Because there was only one variable (different treatment) in the study, the data were subjected to one-way analysis of variance (ANOVA).

score									
	Ν	Mean	Std. Deviation	Std. Error	95% Confidence Internal for Mean		Minimum	Maximum	
					LowerBound	Upper Bound			
1.00	30	3.9333	1.25762	.22961	3.4637	4.4029	2.00	6.00	
2.00	30	3.9667	1.47352	.26903	3.4164	4.5169	2.00	6.00	
3.00	30	3.9000	1.21343	.22154	3.4469	4.3531	2.00	6.00	
Total	90	3.9333	1.30513	.13757	3.6600	4.2067	2.00	6.00	

FIGURE 1: DESCRIPTIVE SCORE OF PRETEST Descriptions

Figure 1 shows the descriptive scores of the pretest, and it is obvious that the three groups demonstrated almost the same result in this pretest.

FIGURE 2: DESCRIPTIVE SCORES OF THE IMMEDIATE POSTTEST Descriptives

Score									
	Ν	Mean	Std. Deviation	Std. Error	95% Confidence Internal for Mean		Minimum	Maximum	
					LowerBound	Upper Bound			
1.00	30	7.5000	1.35824	.24798	6.9928	8.0072	4.00	10.00	
2.00	30	6.3333	2.21800	.40495	5.5051	7.1615	2.00	10.00	
3.00	30	5.2333	2.26949	.41435	4.3859	6.0808	2.00	10.00	
Total	90	6.3556	2.17912	.22970	5.8991	6.8120	2.00	10.00	

Figure 2 displays the means & standard deviations for each group's immediate posttest (raw scores of the multiple-choice recognition test). It shows that in the multiple-choice recognition test, EGI (mean=7.5000) performed better than EG2 (mean=6.3333), and EG2 gained better result than CG (mean--5.2333).

Comparing the mean score of each group between pretest and immediate posttest, we find that, for EG1, the mean score of pretest is 3.9333 and that of immediate posttest is 7.5000; for EG2, the mean score of pretest is 3.9667 and that of immediate posttest is 6.3333; for CQ the mean score of pretest is 3.9000 and that of immediate posttest is 5.2333. The comparison of mean score indicates that for the multiple-choice recognition test, the gain of EG1 is larger than either the EG2 or CG: The gain of EG2 is larger than that of the CG

FIGURE 3:
COMPARISON BETWEEN GROUPS ON THE IMMEDIATE TEST
ANOVA

Scole								
	Sum of Squares	df	Mean Square	F	Sig			
Between Groups Within	77.089	2	38.544	9.705	.000			
Groups Total	345.533	87	3.972					
_	422.622	89						

To make statistical comparisons between groups, the data from the immediate posttest are submitted to one-way ANOVA with treatment as the between-groups factor. For the multiple-choice recognition test, the ANOVA indicated that F=9.705, and p value is .000 to three decimal spaces, less than .05 or. 01, and thus the mean score of the three groups in the immediate posttest is significantly different between groups.

FIGURE 4:
POST HOC TEST
Multiple Comparisons

D	ependent Variable	: Score					
L	SD						
	(I)实验方法 (J)等	实验方法	Mean	Std Error	Sig	95% Confidence Interval	
			Difference(I-J)			Lower Bound	Upper Bound
	1.00	2.00	1.1667*	.51456	.026	.1439	2.1894
	3.00		2.2667*	.51456	.000	1.2439	3.2894
	2.00	1.00	-1.1667*	.51456	.026	-2.1894	1439
	3.00		1.1000*	.51456	.035	.0772	2.1228
	3.00	1.00	-2.2667*	.51456	.000	-3.2894	-1.2439
	2.00		-1.1000*	.51456	.035	-2.1228	0772
	1.00 3.00 2.00 3.00 3.00 2.00	2.00 1.00 1.00	Difference(I-J) 1.1667* 2.2667* -1.1667* 1.1000* -2.2667* -1.1000*	.51456 .51456 .51456 .51456 .51456 .51456	.026 .000 .026 .035 .000 .035	1439 1.2439 -2.1894 .0772 -3.2894 -2.1228	Upper Bound 2.1894 3.2894 1439 2.1228 -1.2439 0772

*. The mean difference is significant at the .05 level.

Figure 4 shows the data analysis of the immediate posttest using Post Hoc Tests which is to test for differences

between all possible pairs of means. The mean score in the immediate posttest between EG1 and CG is significantly different. (p=0.000<0.01). The mean scores between EG2 and EGI; EG2 and CG also differ from each other, but p is 0.026 and 0.035 respectively, less than 0.05, but not less than 0.01.

FIGURE 5:
DESCRIPTIVE SCORE FOR DELAYED POSTTEST
Descriptives

	Ν	Mean	Std. Deviation	Std. Error	95% Confidence Internal for Mean		Minimum	Maximum
					LowerBound	Upper Bound		
1.00	30	6.9667	1.29943	.23724	6.4815	7.4519	5.00	10.00
2.00	30	6.7667	1.07265	.19584	6.3661	7.1672	5.00	9.00
3.00	30	5.5000	1.25258	.22869	5.0323	5.9677	3.00	9.00
Total	90	6.4111	1.36484	.14387	6.1253	6.6970	3.00	10.00

Figure 5 displays the means and deviations for each group's delayed posttest. Two EGs seem to manifest almost the same performance (mean score of EG1 is 6.9667, and that of EG2 is 6.7667), but they still did better than CG (mean=5.5000). The mean score of EG1 in the delayed posttest decreases compared to the score of the immediate posttest (mean=7.5000), however the mean score of EG2 and CG increases more or less in the delayed posttest, since the mean score of EG2 and CG in the immediate posttest is 6.3333 and 5.2333 respectively.

FIGURE 6: COMPARISON BETWEEN GROUPS ON THE DELAYED POSTTEST ANOVA

	Sum of Squares	df	Mean Square	F	Sig
Between Groups Within	37.956	2	18.978	12.916	.000
Groups Total	127.833	87	1.469		
	165.789	89			

Likewise, to make statistical comparisons between groups, the data from the delayed posttest are submitted to one-way ANOVA, with treatment as the between groups factor. For this test, the ANOVA indicates that F=12.916, and p value is still, .000 to three decimal spaces, less than .05 or: 0l, and thus the mean scores of delayed posttest are significantly different between groups.

FIGURE 7:
POST HOE TESTS
Multiple Comparisons

Dependent V	Dependent Variable:Score									
(I)实验方法(J)实验方法		Mean	Std Error	Sig	95% Confidence Interval					
		Difference(I-J)			Lower Bound	Upper Bound				
1.00	2.00	.2000	.31298	.799	.5463	.9463				
3.00		1.4667*	.31298	.000	.7204	2.2130				
2.00	1.00	2000	.31298	.799	9463	.5463				
3.00		1.2667*	.31298	.035	.5204	2.0130				
3.00	1.00	-1.4667*	.31298	.000	-2.2130	7204				
2.00		-1.2667*	.31298	.000	-2.0130	5204				

*. The mean difference is significant at the .05 level.

Figure 7 shows the investigation of the Post Hoc Tests, which is more precise to test the differences between all possible pairs of group for means. From the above, there is no significant difference between EG1 and EG2 because the p value is .799, more than .05. It means in the delayed posttest, EGI did not perform much better than EG2. However, both the EGs still outperformed CQ which we can see from figure 7 (.000<.05 or .01). Figure 7 still places EGI and EG2 in the same subset, whose means are not significantly different .from each other. CG is in a different subset, indicating that its mean does differ significantly from the means of the two EGs.

Finally, the overall results of the retrospective interview are briefly reported here. To the first question, "what did you pay attention to most when you read the input material containing the past hypothetical conditional in English", we found that 56% of participants paid attention to the bolding part of the input sheet; 20% of the participants paid attention to the examples of the input knowledge, and 24% of participants focused on the unfamiliar words and phrases of the input.

To the second question of whether the participants of 2EGs noticed the past hypothetical conditional during reading the material, we found that 80% of the participants in EGI paid attention to the target form while only 60% participants in EG2 noticed that target form embedded in the reading material because they said that when they did the true-or-false questions, it seemed that only some of the answers had some concern with the target form.

To the question about their perception of difficulty during EG1's production attempts, 60% of the participants reported that they noticed that their retelling was somewhat different from the target form, but they were unable to reproduce it because of some unfamiliar words, phrases or working memory capability. And 20% of EGI participants reported that they had difficulty in remembering the meaning or vocabulary, and 60% of EGI had difficulty in

Score

remembering the target form.

To the question that in the immediate posttest, how you chose the answers-by paying attention to the input material (explicit input sheet and reading passage) or by paying attention to retelling the passage (for EG1) / answering the true-or-false comprehension questions (for EG2), 60% participants of EG1 reported that they benefited a lot from the retelling task and 80% of the participants of EG2 reported that the input material did help a lot when they completed the immediate posttest.

To the last question, 20% of the EGI participants and 20% EG2 could not decide whether their choices in the delayed posttest resulted from the input material or the following treatment, because the time interval was too long between the two posttests. 60% of the EG1 still believed that retelling task helped a lot to promote noticing of the target form and 40% of the EG2 participants reported that the input material impressed them deeply when doing the delayed posttest.

V. DISCUSSION AND CONCLUSION



Figure8: Means plot of the three groups in the immediate posttest



Figure 9 Means plot of three groups in the delayed posttest

To summarize the findings of this study in terms of the two research hypotheses, our results confirmed Hypothesis 1, which predicted the output in the form of retelling would promote the noticing of the formal elements in the target form input. From Figure 4, we see that EGI which received the output treatment performed significantly better than both EG2 and CG which received non-output treatment in the immediate posttest. In the delayed posttest, EGI still did better than CG in noticing the past hypothetical conditional in English.

However, this study only partially supported Hypothesis 2, which predicted that EG1 would perform better than EG2 in the noticing of the past hypothetical conditional. From Chapter 4, we know that in the immediate posttest, the mean score between EG1 and CG was significantly different. (p=0.000<0.01), and the mean score of EG1 was significantly higher than that of EG2. It might indicate that EG1 performed much better than EG2 in noticing the target form (see figure 8). However, in the delayed posttest, the two EGs seemed to manifest almost the same performance (see figure 9). There was no significant difference in mean score between these two EGs (see figure 7), which was contrary to our initial expectation.

These two results of the study suggested that retelling task as an output treatment for the EGI did not necessarily trigger the noticing of the target form better than answering true-or-false comprehension questions treatment delivered to the EG2, because there was no significant difference found between the two groups in the delayed posttest, in spite of significant difference in the immediate posttest. Therefore, it cannot be said that retelling as an output is necessarily better at promoting noticing of the target form than answering true-or-false comprehension questions.

Several possible reasons could be considered to account for the results. One was that the treatment period was too short for its effect to manifest itself. In Izumi (2002) study, the output group was given five tasks, containing a total of 46 instances of the target form embedded in them, over a two-week period. In contrast, the present study used two tasks,

with 8 instances of the target form imbedded in the explicit input material and the reading passage, which were given to the participants in less than 3 days. So the duration and the amount of the target form items imbedded in the explicit input material and the reading passage may have contributed to the differences in the results.

In addition, whether the retelling task provided alone could bring about measurable noticing and learning is still an empirical question that needs testing. From the interview, we know that retelling task that targets a specific grammatical structure may promote noticing of the gap, because retelling task maximizes the similarities between the participants' production and the target language model. However, retelling task may often lead different participants to pay attention to different aspects of the reading materiel, such as vocabulary, discourse markers, grammar forms, rhetorical organization, and propositional contents. In other words, if the participants are to complete retelling tasks, they almost need to pay attention to all the aspects of the material. So noticing of the target form may not be a first priority when they do the retelling task.

The true-or-false comprehension task, on the other hand, may have primed the EG2 participants to pay closer attention to the target form. In the language testing literature, comprehension questions have been found to serve as an additional knowledge source that develops the test takers' mental model or meaning construction (Gordon & Hanauer, 1995). The true-or-false comprehension question format used in this study, with its relatively narrow focus and limited options for answering, might have inadvertently caused the EG2 participants to pay greater attention to the target form. For example; to decide whether the statement "At last the farmer planted the tree in a wood. He didn't plant it near the road because he was afraid that strangers would steal the apples " was true or false, the participants had to know that the following sentence from the passage was counterfactural: "If I had planted the tree near the road, strangers would have stolen the apples ". Working with statements such as this one might have promoted more careful attention to and processing of the conditional form in the reading material. Further studies addressing the noticing and learning issues, however, need to be especially careful to prevent inadvertent priming of the comparison group from obscuring possible between-group differences.

Although, compared to answering the true-or-false comprehension questions, the unique effects of the output in the form of retelling in promoting noticing of the target form were not confirmed in the delayed posttest of this study, it is reasonable to assume that producing output might have enhanced the learners' awareness of the inadequacies in their interlanguage performance. In fact, in the immediate posttest, EG1 participants had showed significant improvement on noticing of the target form. However, EG1 were under the heavy cognitive demand of remembering and retelling the given passage accurately in the phase 2 task. It is likely that many of the EGI participants could store the target form in their short-term memory, so EGI performed better and showed significant gain on noticing of the target form in the immediate posttest. However, after a period of time, the high memory load demanded by the retelling tasks was probably too taxing for these EGI participants to conduct careful analysis of the target form, so in the delayed posttest, EG1 failed to show greater gain in noticing of the target form than EG2. To some extent, the retelling requirement of this treatment may have blocked further processing of the target form. If this explanation turned out to be a valid one, it would point to the importance of considering, not only how task facilitates noticing of the form, but also how further processing of the form would be encouraged.

So EGI need to be under relatively low cognitive demand-to retell a relatively simple passage without complicated sentences, unfamiliar words or phrases, but with target forms embedded in it. This is aimed to prevent participants from deviating their attention to other aspects of the passage. If so, it is quite easy for participants to decide which form or which part they need to pay much attention to or notice.

As for true-or-false comprehension questions, in traditional sense, they are not based on form, so learner may just make guesses about the reading material based on the semantic clues to answer the questions even though they encounter the target form that they are not quite clear about. However, in this study, some of the true-or-false comprehension questions were devised as form-based rather than meaning-based, and out of our expectation, these questions have promoted participants of EG2 to pay much attention to the grammatical structure. Therefore, the EG2's performance in the two posttests in this study have revealed that true-or-false comprehension questions, if devised appropriately (i.e., with a focus on form), can also contribute to promoting learners' noticing of the target form.

In addition, one result was also out of our expectation. From figure 1, figure 2 and figure 5, we see that there was gradually an increasing trend in the mean score of CG. This might have just justified, to some extent, the effect of input enhancement on noticing the target form because the treatments that CG received were just studying the explicit input material with the target structure enhanced with bolding and underlining, and then reading a passage. Therefore, input enhancement might still play a role in drawing learners' attention to the target form.

Therefore from the results of this study, it suggests that output does have significant impact on L2 learning, but does not always succeed in drawing the learners' attention to the target grammatical form. Therefore, before conducting an experimental study, researchers and teachers need to make enough preparation and get to know about how to adequately control for learners' otherwise scattered attention if the aim of the research or teaching is to draw learners' attention to a specific form and help them to acquire it.

As for the second research hypothesis, in the present study, only retelling task for EGlwas employed as the output treatment, and we expected that retelling task would prompt more noticing of the target form than answering true-or-false comprehension questions. However, certain factors concerning the retelling task and answering

true-or-false comprehension questions were not well controlled, so some unexpected experimental results emerged.

References

- [1] Celce-Murcia, M., & Larsen-Freeman, D. (1983). The grammar book. Rowley, MA: Newbury House.
- [2] Gass, S. (1988). Integrating research areas: a framework for second language studies. Applied Linguistics, 9, 198-217.
- [3] Gordon, C., & Hanauer, D. (1995). The interaction between task and meaning construction in EFL reading comprehension tests. *TESOL Quarterly*, 29, 299-324.
- [4] Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experiment study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541-577.
- [5] Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-278.
- [6] Izumi, S., Bigelow, M., Fujiwara, & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421-452.
- [7] Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. Applied Linguistics, 10, 52-79.
- [8] Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. Language Learning, 45, 283-331.
- [9] Rutherford, W. (1987). Second language grammar; Learning and teaching. London: Longman:
- [10] Schmidt, R. (1990). The role of consciousness in second language learning. Applied Linguistics, 11, 206-226.
- [11] Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawai'i Press.
- [12] Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*, (pp.237-326). Rowley, MA: Newbury House.
- [13] Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. Applied Linguistics, 2, 159-169.
- [14] Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50, 158-164.
- [15] Swain, M. (1995). Three functions of output in second language learning. In C1 Cook, & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.
- [16] Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 65-81). New York: Cambridge University Press.
- [17] Truscott, J. (1998). Noticing in second language acquisition: A critical review. Second Language Research, 14, 103-135.

Weiwei Pan was born in Jinzhou, China in 1980. She received her Master degree in linguistics from Beijing Foreign Studies University (BFSU), China in 2005, and now she is a PHD student in National Research Center for Foreign Language Education in BFSU.

She is currently a lecturer in Teachers' College of Beijing Union University. Her books and published articles are listed as follows: Zhang Dongchang, Pan Weiwei. 2010. *Empirical Researches on Reading Strategies in English*. University of International Business and Economics Press. Pan Weiwei. 2011. On Cooperative Principle from View of Marriage in *Pride and Prejudice. Journal of Time Literature.* Pan Weiwei. 2010. A Comparative Study of Concealment Function in Chinese and English Euphemism. *Journal of Changsha University*.

Her research interests include Second Language Acquisition and Sociolinguistics.