The Effect of Output Task on EFL Writing: How Far Can It Go?

Hui Huang
Monash University, Australia

Abstract—**The present study attempts to further the understanding of the effect of output in SLA by investigating to what extent an output task (i.e., reconstruction task) is effective in EFL writing. Data in the form of pre-, post- and delayed post-tests were collected from 28 students (14 in input-oriented group while 14 in output group) at one Chinese University. The results revealed a positive and enduring effect of the output task on learners' grammatical accuracy in writing as compared with input-oriented instruction that has long been prevalent in Chinese EFL classroom. However the results did not find an equal impact of the task on different rules: The rules carrying meaning (e.g., connective) or having semantic equivalents in target language (e.g. passive "-ed" or comparative "-er" or superlative "-est") are produced more accurately by the output group while less meaningful morphology (e.g. plural noun "-s", third singular "-s") or local rule (e.g. prepositions) were most difficult for these EFL learners. We argue semantic and functional awareness is the key format for learners to acquire and internalise the grammatical forms.**

Index Terms—**output task, EFL, grammatical accuracy, writing**

I. Introduction

With the proposal of Swain's *Comprehensible Output Hypothesis* (1985, 1995, 2005), output has been viewed not merely as an end product of learning, but as an important factor to promote second language (L2) learning for many L2 teachers and researchers. Many studies, for example, have attempted to compare the empirical effect of output tasks with non-output conditions, such as with input enhancement (e.g., Izumi, 2002, Song and Suh, 2008; Song, 2010) or with processing instruction (e.g., Morgan-Short & Bowen, 2006). These studies have found that output tasks are beneficial to L2 language learning even though variations of such effect exist under different task conditions, e.g. comprehension or production exercise (Dekeyser & Sokalski, 1996). Most of the studies, however, resonated with each other on the positive effect of output tasks on production tasks in English learning context (e.g. Izumi et al, 1999, Izumi, 2002, Song and Suh, 2008) if not all on comprehension tasks. What is not clear is the extent to which output task can be beneficial to L2 production tasks? For example, are output tasks equally effective for the different language rules and how durable such effect is. The answer to this question becomes more complex when different learner profiles are taken into consideration. This article intends to explore some answers to these issues within Chinese teaching English as a foreign language (EFL hereafter) setting, particularly in an attempt to examine to what extent one frequently-used output task (i.e., reconstruction task) will benefit student's production skills in a natural writing context (i.e., writing short essays).

II. Literature Review

A. EFL Background in China

In spite of reforms in English education over decades in China (e.g., the introduction of communicative language teaching since late 1980's or task-based language teaching in 1990's), teaching EFL in China is still test-driven, mainly aimed at preparing learners for all kinds of national or graduation tests. In practice, there are some attempts for communicative or task-based language teaching, however most EFL classrooms in China are still featured by a preoccupation with careful, often painstaking examinations of grammar or vocabulary items which are always the focus of many tests (e.g., Rao, 2013). As a result grammatical accuracy is much more emphasised in practice as compared with fluency even though most English teachers agree to the fact that accuracy and fluency are both essential in language teaching and learning (Rao & Lei, 2014). This is, at least partially attributed to the lack of teaching staff. The EFL classroom in China is so big (always more than 50 students) that only a teacher-centred approach is welcome and assumed to be effective and communicative language teaching methods are not as effective as expected, either due to the students' negative attitudes (e.g., Rao, 2002) or to the teacher's dilemma of teaching grammar (Ouyang, 2000; Zheng and Borg, 2014). To cater for the national examinations, many highly controlled exercises used in Chinese EFL teaching, the majority of which are in the form of multiple-choice comprehension questions. In the situation described above, students' responses in spontaneous language production performance (e.g., writing an essay) leave many teachers feeling very disappointed at their students' inability to use what they have been repeatedly taught. For example, students often make such mistakes as "once a time, there is…" or "I'm go to…". This frequently occurs despite the fact that most students are already very skilful in choosing satisfactory answers on multiple-choice grammar tests.

B. Output Tasks and EFL Grammar Teaching

TABLE 1
A SUMMARY OF EFL STUDY USING OUTPUT TASKS

EFL study using output tasks	Methodology	Target features	Results
Izumi (2002): to compare the facilitative effects of output and visual input enhancement on the acquisition of target form	1. Subjects: Five groups (61 college-level students from two universities): 11 in +O+IE group, 12 in +O-IE group, 12 in -O+IE group, 12 in -O-IE group and 14 in CG. The subjects have diverse L1 background; 2. Pre- and immediate post-test; 3. Tests: sentence combination test, picture-cued completion test, interpretation test and grammaticality judgment test;	Relativisation	1. Those in output tasks outperformed those in the input-enhancement tasks; 2. Those in input enhancement tasks failed to show measurable gains in learning;
Izumi et al. (1999): To test the role of output on noticing and performance on the target form	1. Subjects: Two groups (22 college-level students from an ESL department with diverse L1 background):11 in EG and 11 in CG; 2. Pre- and two post-tests which were conducted immediately & the day after the two respective treatments; 3. Grammaticality Judgment test and picture-cued production test;	Past hypothetical conditional	1. Phase 1 tasks resulted in noticing and immediate incorporation of the target form, the posttest performance failed to reveal the effects; 2. Phase 2 tasks resulted in improvement on post-test 2.
Song & Suh (2008): To test the role of output and two types of output tasks (reconstruction and picture-cued writing tasks)	1. Subjects: Three groups (52 intermediate Korean EFL students with some partial knowledge of the target form in three groups: EG1, EG2 and CG; 2. Pre- and immediate post-test; 3. Recognition and written production tests;	Past counterfactual conditional	1. Noticing occurred more in the output groups as compared with non-output conditions 2. Participants with output treatments performed significantly better than those in the non-output condition on the production post-test
Song (2010): To examine the effectiveness of an output practice (i.e., translation) on promoting noticing and acquisition of lexical phrases	1. 36 EFL students from two parallel classes of same level, divided into EG (18) and CG (18); 2. Pre- and immediate post-test; 3. Multiple-choice questions and Chinese-English translation;	Lexical phrases	1. EG noticed the target form more than the CG; 2. The effect of output practice on promoting acquisition of lexical phrases was much greater than that of input practice
Abadikhah and Zarrabi (2011): To examine whether engaging learners in output tasks can fill the gap between comprehension and production of verbal morphemes	1. Iranian EFL learners (20 output group and 18 control group); 2. Pre- and post-test;	Verbal Morphemes (-ing and –ed)	1. Output group outperformed the control group 2. Control group failed to show comparable improvement in their production although they had measurable gains in comprehending the target linguistic form
Abadikhah (2012): To examine the effectiveness of mechanical and meaningful production of output on the learning of English relative clauses by Iranian EFL learners.	1. 36 Iranian EFL learners in two groups; 2. Pre- and post-test; 3. Recording of collaborative dialogue;	Relative clauses	1. Both groups improved significantly 2. More meaningful activities elicited significantly more language-related episodes compared to the more mechanical activities
Khatib, M., & Alizadeh, M. (2012): To examine the effects of using two different types of output tasks on noticing and learning the English past tense	1. 60 female subjects in three groups (picture-cued writing tasks, reconstruction tasks & comprehension check-up task; 2. Pre- and post-test;	Past tense	1. Only the reconstruction group improved in their noticing of the target feature. 2. Both picture-cued and reconstruction groups equally promoted their learning of the form.

The *Comprehensible Output Hypothesis* (Swain, 1985, 1995, 2005) claims that producing output in L2, as opposed to merely comprehending the language, may force the learner to move from semantic processing to syntactic processing. According to the hypothesis, producing the target language (TL) may serve as "the trigger that forces that learner to pay attention to the means of expression needed in order to successfully convey his or her own intended meaning" (Swain,

1985, p.249). That is, producing the target language provides learners with unique opportunities for a level of processing (i.e., syntactic processing) that may be needed for the development of target-like proficiency or the enhancement of accuracy (Swain and Lapkin, 1995). In addition, it is also hypothesised that output can promote language acquisition by making learners notice problems in their interlanguage (IL) (Swain and Laptin, 1995) and thus prompting learners to do something about those problems. For example, they can seek out relevant input with more focused attention, or search for alternative means to express the given intention and stretch their IL capability, formulate and test a hypothesis, and modify it upon receiving feedback (Izumi et. al., 1999).

Much recent empirical EFL research has evidenced the effect of output conditions on learning target forms in production exercises (see TABLE 1 for details). Izumi and his associates (e.g., Izumi, 2002; Izumi and Bigelow, 2000; Izumi et al., 1999), for example, have conducted a series of ESL studies to compare output asks with non-output tasks. Their results are mixed, two of which are in favour of the beneficial effects for output groups. Izumi (2002) has attributed the superior performance of output groups to the effect of depth of processing and integrative processing. Another strand of research on EFL using output tasks is focused on investigating the effect of different output tasks on ESL learning. For instance, a study with the EFL Korean students (Song and Suh, 2008) has indicated that the participants receiving two various output opportunities (i.e., reconstruction and picture-cued writing tasks) performed significantly better than those in the non-output conditions on the production post-test even though there were no significant difference between the two different methods. The same effect of output tasks on target forms is also found in other Asian EFL contexts (e.g. Song, 2010 in China, Abadikhah and Zarrabi, 2011; Abadikhah, 2012 in Iranian context)

All these studies measure their subjects' learning of the targeted forms by using contextualised production tasks (e.g. picture-cued writing in Izumi's studies; sentence completion task in Song & Suh or in translation work Song, 2010). None of them, however, have examined the subject's English grammatical accuracy in natural language setting, like writing an essay. When natural writing is used as assessment task, it is very hard for learners to pay as much attention to accuracy of grammatical forms when most of their attention is directed to semantic processing of sentences or idea development (VanPatten, 1990; Huang, 2008).

Additionally, among these studies one issue that has received relatively less attention from researchers is the question: to what rules can output tasks be effectively applied? The many studies on output tasks so far are only targeted at one or two grammatical features and most of them are syntactical forms. In TABLE 1 we can see the tested forms in the series of EFL studies were either the relative clause of English or English past hypothetical conditional or the lexical phrases. Even though the relativization and English past hypothetical conditional are complex in syntax, processing the output tasks (through either reconstruction or picture-cued writing tasks), however, means a necessity to process the target syntactical rules of relativisation or past conditional because the meaning of the rules have been embedded in the sentence. That is what is called by Loschky (1994) as "task essentialness" (i.e., in order to complete the tasks, the meaning of the whole sentence needed to be processed, including the target forms). However other less salient forms especially those morphological forms such as third person singular "-s" or plural pronoun "-s" has very rarely been examined in the vein of the effect of output except for Abadikhah and Zarrabi (2011) who examined two verbal morphemes "–ing" and "–ed" (for details of the study see TABLE 1). As Muranoi (2007) posited whether output practice leads to L2 development heavily depends on various factors one of which he identifies as the linguistic features of the target form.

Another research gap of using output tasks in EFL setting is long term effect: Except for Izumi et al. (1999), the other studies only look at the immediate effect of the output tasks. However the question of the long-term effects of treatment is a very important issue. If the effects of intervention are short-lived, then as some have claimed (e.g. Krashen, 1985), acquisition results from exposure to input and nothing else. If long-term effects can be found, then such a position is less tenable (VanPatten and Fernandez, 2003).

This study, therefore, in an attempt to fill in these research gaps, is intended to examine for what rules output tasks are useful in spontaneous essay writing context in order to further the understanding of the effect of output tasks in EFL, in particular by exploring answers to the following questions:

1) Compared with input-oriented group, do reconstruction output task have a positive effect on their performance accuracy in natural writing context?

2) If such effect is found, will it be durable?

3) For what rules is the output task most / least effective?

III. THE STUDY

This study was an quasi-experimental study conducted in two classes using a design of pre-, post- and delayed post-test with post-test examining the immediate effect while delayed post-test the long-term effect.

A. Subjects

The subjects of this investigation were 28 Chinese University EFL students from two existing classes. Based on their own class, they were identified as two groups: the output group (OG) and the input-oriented group[1] (IG) (each group

---

[1] The group did some controlled exercises over the experimental period. However, there were no any substantial output opportunities in the group.

involved 14 participants). The data from a background questionnaire showed that the participants in both groups were reasonably equivalent in their English learning background, having studied English for 7 years according to the national syllabus.

TABLE 2
PROFILE OF THE SUBJECTS

	OG	IG	Total
No. of Males	7	6	14
No. of Females	7	8	14
Total Population	14	14	28
Average Age	20.8	20.3	
Min. Years of Learning English	7	7	

B. Data Collection and Procedures

The data collection of the study lasted 16 weeks over three stages in two Chinese EFL classroom settings. At the pre-testing stage, a questionnaire on the subjects' background was collected. Then students were asked to complete a 20-minute pre-test in the classroom under the supervision of teachers. The first data from pre-test results was taken as the baseline referential standard to examine the subjects' progress. During the 12-week experiment stage, those students in the OG continuously participated in reconstruction tasks in normal classes whereas the IG received formal, input-oriented instruction on the grammatical features without these activities. The teacher in the IG group taught the grammatical features in a way of input-oriented instruction (that is, giving presentations and examples of the target grammatical features which is followed by the controlled grammar exercises). At the conclusion of the experiment period, all subjects were tested immediately and again four weeks later in order to determine the post-experimental and delayed post-experimental performance on grammatical knowledge. The administration procedures of the post-testing and the delayed post-testing measures were exactly the same as those in the pre-test.

C. Development of Materials

1. Selection of grammar structures

Before the actual study, a preliminary study was conducted with another group of Chinese students at the same university to obtain information on the frequency of grammatical errors that students made in their writing and grammar tests. Then the researcher selected the most frequent errors and classified them into six groups[2] as the focus of the grammatical instruction for both classes in the research. They were: connectives, third person singular verb"-s", plural nouns, prepositions, non-infinite verbs (that is English morphemes of "–ing" and "–ed") and counterfactual conditional[3].

2. Development of tasks and materials

In the study, reconstruction activities were selected as the output task for the experiment. This type of activity is basically meaning-based communicative activity. The reconstruction procedure used in the study was known as a "propositional cluster" task (Rutherford, 1987). In this, students were presented with content words and instructed to reconstruct the text by inserting appropriate function words, linking words, inflectional morphemes, and/or changing word order to produce an accurate, meaningful, and appropriate text. The detailed procedure included: 1) Content words in English were presented with a topic to each student; 2) Students worked on the task collaboratively, in pairs or small groups, producing one complete Chinese version in oral form per group; 3) Each student reconstructed a version of the text from the content words based on the discussed Chinese version; 4) The completed version was then used in subsequent class work: Students were told to compare their version firstly with others and then with the original text; 5)The various versions were analysed and compared and learners refined their own texts in the light of the shared scrutiny and discussion. The efficacy of the output reconstruction activity has been researched in previous studies (e.g., Izumi, 2002).

The texts used in the reconstruction activity were carefully selected and different ones were used for different grammar rules. In each text, the focused grammar structures were operationised in a short, coherent 80-word passage. Each week one text was used, which focused on one linguistic feature and every two weeks the linguistic focus was changed, unnoticed by students.

In class tests, students were required to write a 120-word long passage based on a given title. The title in each test differed each time to avoid that students would recite the composition after the pre-test. However, all essay titles were of the same genre (respectively titled as "My View on Children's Heavy Bags", "My View on Making Friends" and "My View on Examinations"), which required the students to express views on a certain issue.

---

[2] The researcher identified and grouped the grammatical rules on the basis of the grammar book prevalently used in China, such as 《薄冰高级英语语法》 (Bo Bing Advanced English Grammar), 《实用英语语法 》 (*Practical English Grammar*).
[3] Or "subjunctive mood" in some grammar books in China.

D. Scoring and Data Analysis

In scoring the short essay, all errors were marked as minus points first based on their frequency of occurrence. All of these errors were classified into various categories as erroneous / absence of nouns, verbs, and others (see TABLE 6 for details, the highlighted columns being the target rules). The participants in each group were requested to write about 120 words for each text, however, some wrote more than 200 words while some only reached 70 words. Thus it was hard to say whether student A who committed 11 errors in his 240-word text performed worse than student B who only made 5 errors in a 70-word text. Considering greater accuracy, we recalculated the errors of each text based on 120-words rather than its actual word count. For the sake of data processing, these minus marks were recalculated against a full mark of 30. Take the student A as an example. His errors were recalculated and scored as -5.5 and his essay mark was scored as 24.5 in the end. In order to enhance the objectivity and reliability of the data, two experienced teachers worked separately to score all essays.

In order to examine the within-in group progress over time, paired T-tests were used while one-way ANOVA was processed to explore the between-group significance of differences in details.

IV. RESULT

The scores of three essays for the IG and the OG are presented in TABLE 3.

TABLE 3
DESCRIPTIVE OF TWO GROUPS' ESSAY SCORES

	IG (N=14)		OG (N=14)	
	Mean	SD	Mean	SD
Pre-test	18.42	5.56	18.42	4.57
Post-test	17.48	5.38	20.64	4.52
Delayed post-test	19.85	3.85	23.35	2.76

In order to answer the questions addressed earlier, the learners' grammatical performance in essays was compared in details on the basis of: (a) proficiency between groups and (b) proficiency within a group.

A. Between-group Analysis

TABLE 4
BETWEEN-GROUP COMPARISON USING ANOVA

Pre-test	Post-test	Delayed post-test
$F_{(1,26)}=000$, $p=.998$ $Eta^2= .00$	$F_{(1,26)}=2.83$, $p=.104$ $Eta^2= .10$	$F_{(1,26)}=7.62$, $p=.010$ $Eta^2= .23$

In the pre-test, both groups performed very similarly in the two sections ($p>.05$). Therefore the between-group difference after the experimental period was still processed using paired t-test. The results found that the OG, as compared with the IG, achieved significant gains after the experiment and the significance was maintained in the delayed post-test 4 weeks later (TABLE 4) with medium to large effect size (eta^2 = .10 and .23 respectively[4]). Thus, the reconstruction output task adopted in the OG appeared to have functioned more effectively than the input-oriented grammar lessons in spontaneous writing and in addition this level of proficiency had been maintained 4 weeks later.

B. Within-group Analysis

Besides the between-group analysis, analyses were also conducted within each group to examine to what extent the respective tasks had an impact on students' performance of grammatical accuracy. The within-group analysis found that in the spontaneous writing task, the OG group improved significantly over the time while the IG's performance kept very similar across the three tests. This reveals that the output task worked very well in improving students' grammatical performance in writing.

TABLE 5
WITHIN-GROUP RESULTS USING PAIRED T-TEST

	Post- & Pre-test	Delayed Post- & Pre-test
IG	Paired t = -0.85, $p>.05$, Eta^2= .05	Paired t = 0.92, $p>.05$, Eta^2= .06
OG	Paired t = 2.22, $p<.05$, Eta^2= .27	Paired t = 5.34, $p<.05$, Eta^2= .69

In order to understand clearly on what rules this reconstruction tasks have impact, all the mistakes in the essays were carefully scrutinised and categotised (see TABLE 6 for the details). An in-depth analysis was conducted to examine the OG's improvement/maintenance in accurately using different grammar forms.

---

[4] Cohen (1988)，eta^2 > .01 means small effect; > .o6 moderate effect; >.14 large effect.

TABLE 6
MISTAKE COUNTS IN THE ESSAYS OF THE OG GROUP

Group	Test	Missing subject	Missing predict	Conjunctives	3rd person sing.-s	Word order	Article	Noun (plural,-s)	Noun (aspect)	Pronoun	Verb.-tense	Verb-passive	Modal v.	Non-finite verbs (-to;-ing &-ed)	Adjective	Adverb	Preposition	others	Total
IG	Pre	7	14	9	5	30	4	13	4	11	4	4	6	16	3	1	24	17	172
	post	7	16	14	6	21	3	18	2	12	7	1	1	25	1	1	28	15	178
	Delayed	6	11	11	14	15	11	13	0	7	17	2	3	8	7	0	21	17	163
OG	Pre	4	7	21	3	17	7	16	9	4	16	5	4	11	8	1	16	14	163
	post	1	8	10	2	13	8	15	0	11	2	1	2	12	2	0	22	13	122
	Delayed	0	12	6	5	10	5	25	1	4	11	0	0	5	1	0	16	13	114

TABLE7.
THE MOST AND LEAST IMPROVED/MAINTAINED RULES IN THE OG
BASED ON CORRECT PERCENTAGE[1]

	Most Improved/maintained Rules			Least Improved/maintained Rules		
	Conj.	Passive-ed	Adj.[4]	Non-finite verbs	Prep.	Noun (-s).
Pre-test(%)	72	88.59	52.94	96.63	89.61	93.1
Post-test(%)	86.84	97.66	80	96. 12	83.71	84.72
Delayed(%)	94.92	96	80	95.42	90.94	93.37
Gains 1[2] (%)	14.84	9.07	27.06	-0.51	-5.9	-8.38
Gains 2[3] (%)	22.92	7.41	27.06	-1.21	1.33	0.27

1. Correct percentage = correct frequency of the form / frequency of occurrence this form in articles *100%
2. Gains 1= correct percentage of post-test – correct percentage of pre-test
3. Gains 2 = correct percentage of delayed post-test –correct percentage of pre-test
4. In the adjective part, it only refers to the comparative and superlative use of adjectives.

TABLE 7 gives a summary of correct percentage of rules in each section across the tests in the output group with target rules highlighted. The rules that were most and least effectively improved in essay were identified based on the percentage gains across the tests. Among the most improved rules, we found that the rule carrying meaning such as conjunctives between clauses and the ones having semantic equivalence in Chinese such as the passive "-ed" of the verbs and the comparative and superlative adjective were consistently improved or maintained by the subjects. Conversely, less meaningful morphology such as non-finite verbs, plural noun "-s" and third person singular verb "-s" and the local errors in preposition are the least effectively corrected or maintained. Interestingly, the passive "-ed" and correct use of comparative or superlative adjectives were not the focus of treatment, but both emerged as the most effective rules.

V. DISCUSSION

The findings obtained from the data in terms of the questions addressed in the study can be summarized as follows: 1) Compared to input-oriented instruction, the output task conducted in the output group played a significant role in promoting Chinese students' grammatical accuracy in natural writings. The delayed post-test 4 weeks later showed a durable effect of such tasks in maintaining that improvement. 2) The output activities had divergent effects on learning different rules. The rules carrying meaning themselves such as conjunctives, the ones having semantic equivalence in Chinese (e.g., passive verb-ed; comparative and superlative "-er" and "-est") were more easily acquired and/or maintained by the students. On the other hand, less meaningful morphology to the learners such as non-finite verbs, number of nouns and local errors such as prepositions, were the most difficult rules for the Chinese learners.

Among the six target rules (i.e., connectives, third person singular verb"-s", plural nouns, prepositions, non-infinite verbs and counterfactual conditional), only conjunctives was among the most improved rules while infinite verbs, plural noun "-s" and preposition were the least. Even though the third person singular "-s" did not appear as one of the least effective rules[5] in OG group, the mistake count of this rule increased in the delayed post-test. It is interesting to see that there was no any mistake in counterfactual conditionals in students' essays. After a careful reading of all essays, we found that no counterfactual conditional was used, which means that these students may have not understood or may have avoided using the form. An alternatively explanation is that these topics don't necessitate the use of this form.

In the study, the morphological salience is of great importance (Skehan, 1996) in learning grammatical rules. The grammatical rules in the form of lexis (i.e., conjunctives or the comparative "more" or superlative "most") in and of itself are much noticeable, which is the prerequisite to acquisition (Schmidt, 1990). In contrast, other rules suffixing a word (i.e., plural noun "-s", third person singular "-s") are not very salient in morphology for these learners to notice especially when writing systems between L2 English and their L1 Chinese are entirely dissimilar. However we suppose what matters in the process of conducting reconstruction tasks might be semantic awareness through explicit negotia-

---

[5] This might be attributed to the low frequency of the use of the rule in the essays.

tion of meaning. The students in OG, in the context of "making meaning" (Swain, 1998) of words, were mainly focused on the negotiation of meaning; however they may also be exposed to "grammaticised" lexis (Rutherford, 1987). The reconstruction task may have enabled the students to explore the relationship between the form of the "grammaticised" words (e.g. conjunctives, passive voice "-ed" or comparative "-er" or superlative "-es") and their functional use. During the process, they might have had to process the meaning of passive "-ed" or comparative "-er" in order to achieve a meaningful text. Namely completing the reconstruction task meant processing the target forms and the meaning of these forms. As compared to those rules (e.g., non-finite verbs and third person singular "-s") which don't have any semantic equivalence in Chinese, these rules convey much meaning to the Chinese learners because there are semantic and functional counterparts in Chinese. Namely, it is the semantic and functional rather than morphological equivalence that is the key format for learners to acquire and internalise the grammatical forms.

In contrast, in the input-oriented situation, the teacher did attempt to draw the students' attention to the target forms by asking them to focus on the target rules and also through their elaborations of the target rules. However, these external agents did not work very helpfully as hoped. On one hand, despite the teachers' attempts to emphasize these forms and attract the learners' attention, not all students followed the teachers. In this case, the students did not necessarily notice the rule, nor did they learn it (Schmidt 1990). On the other hand, even for the students who followed the teachers carefully, these externally-imposed explanations could not guarantee the learner, who is the real agent of the learning process, was fully cognitively engaged in the teachers' elaborations. For one thing, learners may sometimes fake their comprehension knowledge and pass themselves off as having understood (Izumi, 2003). For another, the teacher's explanations were not necessarily based on where the learners were in the learning process (Pienmann, 1989). This supports the view that this input-oriented instruction often does not work, particularly when acquisition is measured in relation to spontaneous activity because target rules are provided in a way that prevents the normal process of acquisition from operating smoothly.

VI. CONCLUSIONS AND RECOMMENDATIONS

The study, in an attempt to explore the extent to which the output has impact on Chinese EFL grammatical accuracy in natural writing task suggests that a better understanding of the role of the output task in instructional techniques will contribute to the development of learners' IL and their communicative accuracy. The findings of this study have implications for the field of EFL and L2 teaching. When designing a L2 task, the most important thing is to optimize every opportunity to fully get the learners cognitively involved in the learning process. The research findings that the output task is not equally effective in linguistic forms indicates that an essential component of a successful pedagogy lies in a careful examination of both the object and the agent of learning that is, the target language and the learners.

However, the results obtained from the study are tentative --- open to challenge and verification because the study suffers from certain limitations. For instance, the number of students was very limited; the assessment of grammatical errors in writing had an element of subjectivity; and the designed grammar tests may have emphasized some factors more than others. The study therefore suggests a few possible directions for future research. More research is clearly desirable into how such tasks can be adapted to take account of individual differences, particularly with more qualitative and descriptive data. Measurement is a very important factor to be considered in assessing L2 attainment (e.g., Ellis, R., 2005). Future research thus needs to explore the extent to which output attempts might affect other controlled and spontaneous situations.

REFERENCES

[1]   Abadikhah, S. (2012). The effect of mechanical and meaningful production of output on learning English relative clauses. System, *40*, 1, 129-143.
[2]   Abadikhah, S. & Zarrabi, F. (2011). The Effect of Output Tasks on the Acquisition of English Verbal Morphemes. *Theory and Practice in Language Studies*, Vol 1, No 11 (2011), 1549-1560.
[3]   Carr, T. And T. Curren. (1994). Cognitive factors in learning about structural sequences: application to syntax'. *Studies in Second Language Acquisition* 16: 205-230.
[4]   Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Earlbaum.
[5]   DeKeyser, R. and K. Sokalski (1996). The differential role of comprehension and production practice. *Language Learning* 46, 613–642.
[6]   Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language, *Studies in Second Language Acquisition*, 27, 141-172.
[7]   Huang, H. (2008). Learner-centred L2 learning through learner elaborations: A cognitive approach. Kunming: Yunnan University Press.
[8]   Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24, 541-577.
[9]   Izumi, S. (2003). Comprehension and production processes in second language learning: in search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24, 168-196.
[10] Izumi, S. & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34(2), 239-278.
[11] Izumi, S., 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: 421-452.

[12] Khatib, M., & Alizadeh, M. (2012). Output Tasks, Noticing, and Learning: Teaching English Past Tense to Iranian EFL Students. *English Language Teaching*, *5*(4), 173-187.

[13] Krashen, S.D. (1985). The Input Hypothesis: Issues and Implications. Harlow, Longman.

[14] Loschky, L. (1994). Comprehensible input and second language acquisition. *Studies in Second Language Acquisition*, 16(3): 303-323.

[15] Morgan-Short, K. & H. Bowen (2006).Processing instruction and meaningful output-based instruction'. *Studies in Second Language Acquisitio*n, 28: 31-65.

[16] Muranoi, H. (2007). 9 output practice in the L2 classroom. in R. Dekeyser (ed.) *Practice in a Second Language: Perspective from Applied Linguistics and Cognitive Psychology*. 51. Cambridge: Cambridge University Press.

[17] Ouyang, H. (2000). One-way ticket: A story of an innovative teacher in Mainland China. *Anthropology and Education Quarterly*, 31(4): 397-425.

[18] Pienmann, M. (1989). Is Language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics* 10: 52-79.

[19] Rao, Z. (2002). Chinese students' *perceptions* of communicative and non-communicative activates in EFL classroom. *System*, 30: 85-105.

[20] Rao, Z. (2013). Teaching English as a foreign language in China: looking back and forward. *English Today*, *29*(03), 34-39.

[21] Rao, Z., & Lei, C. (2014). Teaching English as a foreign language in Chinese universities: The present and future. *English Today*, *30*(04), 40-45.

[22] Rutherford, W. (1987). Second Language Grammar: Learning and Teaching. London: Longman.

[23] Schmidt R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics, 11: 129-158.*

[24] Song, Z. (2010). An Empirical Study of the Role of Output in Promoting the Acquisition of Linguistic Forms. *English Language Teaching*, *3*(4): 109-114.

[25] Song, M. & B. Suh. (2008). The effect of output task types on noticing and learning of the English past counterfactual conditional. *System*, 36: 295-312.

[26] Swain, M. (1985). Communicative competence: some roles of comprehensible and comprehensible output in its development. in: S. Gass & C. Madden (Eds) *Input in Second Language Acquisition.* Rowley, MA, Newbury House.

[27] Swain, M. (1995). Three functions of output in second language learning. in G. Cook and B. Seidhofer (eds.): *Principle and Practice in Applied Linguistics*. 1995. Oxford: Oxford University Press.

[28] Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning*. Mahwah, NJ: Lawrence Erlbaum.

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

[30] VanPatten, B. (1990). 'Attending to content and form in the input: an experiment in consciousness's'. *Studies in Second Language Acquisition*, 12: 287-301.

[31] VanPatten, B. & Fernandez, C. (2003). The long-term effects of processing instruction. in: B. VanPatten (Ed) *Processing Instruction: Theory, Research and Commentary* (Mahwah, N.J. , L. Erlbaum Associates).

[32] Zheng, X., & Borg, S. (2014). Task-based learning and teaching in China: Secondary school teachers' beliefs and practices. *Language Teaching Research*, vol. 18(2): pp. 205-221.

Hui Huang is a lecturer at School of Languages, Literatures, Cultures and Linguistics at Monash University. She obtained her PhD in applied linguistics from La Trobe University, Australia in 2007. She has been researching and publishing in the areas of second language acquisition, language pedagogy and technology, cross-cultural communications and immigrant identity.