The Journal of Applied Linguistics Vol. 2, Issue 2
Fall 2009

The Effect of Comprehensible Input and Comprehensible Output on the Accuracy and Complexity of Iranian EFL Learners' Oral Speech

Omid Tabatabaei[*]
Masumeh Yakhabi
Islamic Azad University, Najafabad Branch

This study aimed at investigating the relative impact of comprehensible input and comprehensible output on the development of grammatical accuracy and syntactic complexity of Iranian EFL learners' oral production. Participants were 60 female EFL learners selected from a whole population pool of 80 based on the standard test of IELTS. To investigate the research questions, the participants were randomly divided into three groups: Input group, output group, and control group. The study involved two phases: the pre-task phase, and the main-task phase. During the pre-task phase, the input group received comprehensible input. In the same phase, the output group was pushed to be engaged in comprehensible output production. The control group neither received input, nor was engaged in output production. In the main-task phase, all subjects performed monologues that were separately recorded, and later transcribed and coded in terms of accuracy and complexity through Bygate's (2001) standard coding system and finally scored. The statistical analysis of the results revealed that while the output group outperformed the input group in grammatical accuracy, the input group proved to be more rigorous and influential in developing speech complexity. The study supports Swain's (1985) claim that there are roles for comprehensible output that are different from and independent of comprehensible input, and Skehan & Foster's (2001) theory

[*] Corresponding author. E-mail: tabatabaeiomid@yahoo.com

regarding human beings' limited attentional capacities that can be devoted to one aspect of oral speech at the expense of the other. Generally, it is implied that the most effective way for improving oral speech, based on the literature and the results obtained from this study, is an eclectic approach which conflates both comprehensible input and comprehensible output.
Keywords: Comprehensible Input, Comprehensible Output, Accuracy, Complexity

The assessment of L2 speaking is a relatively new venture within the young discipline of applied linguistics (Fulcher, 2003). Speaking is a complex construct that is mainly operationalized through the three measures of 'complexity', 'accuracy', and 'fluency'. In other words, the development of speaking ability is measured in terms of progresses made in each of these aspects (Skehan, 1998). According to Finardi (2008), speaking an L2 is an important and universal activity one aspect of which is usually achieved at the expense of others as learners conceptualize, formulate and articulate messages. Skehan (1998) points out that there is a tension between 'complexity', 'accuracy', and 'fluency'; what he calls a 'trade-off' effect. Finardi (2008) investigated whether the repetition of a picture description task will result in gains in complexity, and whether these gains are paid off especially by gains in accuracy. The results of the study corroborate evidence of trade-off effects in oral production, especially in terms of complexity at the expense of accuracy.

Bygate (1999) defines accuracy as the extent to which a speaker's selection of the formal features of the language (vocabulary, idiomatic phrases, grammatical morphemes, pronunciation patterns in speech) correspond to patterns that a representative section of the target population of speakers would find normal, and avoiding what they would find abnormal, for the meanings being conveyed. Complexity, however, reflects the quality of structures used in the communication in terms of the number of words clustered within structures, and the extent to which structures are clustered together.

Regarding what factors can help second language learners to improve their oral proficiency, Comprehensible Input Hypothesis (Krashen, 1981) and Comprehensible Output Hypothesis (Swain, 1985) are two theories each having their own claims regarding the most influential variable (comprehensible input, output) for improving oral speech and its various dimensions.

The Input hypothesis is Krashen's attempt to explain how the learner acquires a second language. According to this hypothesis, the learner improves and progresses along the 'natural order' when s/he receives L2 'input' that is one step beyond his/her current stage of linguistic competence (Schutz, 2007). In the input-based programs, students' grammatical errors are ignored in favor of enhancing the students' fluency. Feedback is considered necessary only when intelligibility is affected. Krashen's fear of attention to accuracy comes from the idea that it may get in the way of fluency and could make the learner self-conscious and as a result, impede second or foreign language acquisition. Accuracy in natural approach is thought to evolve as the learner progresses without conscious attention to form or error correction (Harati, 2000).

Krashen (1982) claims that the function of the L2 learners' output (production) in second language acquisition is only to provide a further source of comprehensible input indirectly by inviting more input from speech partners. Krashen (1981) suggests that real language acquisition develops slowly, and speaking skills emerge significantly later than listening skills, even when conditions are perfect. The best methods are, therefore, those that supply 'comprehensible input' in low anxiety situations, containing messages that students really want to hear. These methods do not force early production in L2, but allow students to produce when they are 'ready', recognizing that improvement comes from supplying communicative and comprehensible input, and not from forcing and correcting production. (1981, pp. 6-7)

However, research findings show that predominantly meaning-focused instruction does not lead to native-like grammatical competency. Evaluation of the immersion programs has shown that while students, who have been provided with comprehensible input acquire considerable fluency and possess

discourse skills, they do not show an equivalent control over a range of grammatical items (Mangubhai, 2001). In fact, Contrary to Krashen's insistence that acquisition is dependent entirely on comprehensible input, most researchers now acknowledge that learner output (Swain, 1985) also plays a part in SLA (Ellis, 2008). In response to the shortcomings and criticisms leveled against comprehensible input hypothesis, Swain (1985) put forward the Comprehensible Output Hypothesis claiming that production can aid a more complete acquisition when the learner is pushed to produce appropriate language that is grammatically accurate.

According to Swain (1995), comprehensible input may not be sufficient for the acquisition of certain aspects of L2 and that comprehensible output may be needed. Based on this hypothesis, learners must also be given the opportunity to produce comprehensible output. Thus, the role of output is to provide opportunities for contextualized and meaningful use of language, to test out hypotheses about the target language, and to move the learner from a purely semantic analysis of the language to its syntactic analysis.

Swain (1985) believes that comprehensible output under certain conditions serves specific functions that can draw the learners' attention to the form of language and helps them to move from semantic to syntactic processing. These functions involve:

- *Noticing*: Swain & Lapkin (1995), considering the "noticing/triggering function" of output argue that under some circumstances, the activity of producing the target language may prompt L2 learners to recognize consciously some of their linguistic problems: it may bring their attention to something they need to discover about their second language (possibly directing their attention to relevant input).
- *Hypothesis Testing*: Producing output is one way of testing one's hypothesis about the target language. Learners can judge the comprehensibility and linguistic well-formedness of their interlanguage utterances against feedback obtained from their interlocutors. (Soleimani, Ketabi, & Talebinejad, 2008).

- *Metalinguistic Awareness*: As Swain (2008) has stated, the claim here is that using language to reflect on language producing by other and the self mediates L2 learning.

However, comprehensible output has its own criticisms. As Shehade (2002) has stated, after over a decade of research on Swain's (1985) comprehensible output (CO) hypothesis, there is still a severe lack of data showing that learner output or output modifications have any effect on L2 learning. As stated by Krashen (1998), Output and especially comprehensible output is too scarce to make a real contribution to linguistic competence. Even when the language acquirer does speak, they rarely make the types of adjustments that the CO hypothesis claims are useful and necessary to acquire new forms. Moreover, there is some evidence that suggests that students do not enjoy being "pushed" to speak. This raises the affective filter and thus hampers acquisition.

According to Foster and Skehan (1996), complexity reflects how learners can use the forms closer to the cutting edge of interlanguage development and is more associated with learners' willingness to take risks to use the language with which they are not familiar.

However, as review of literature reveals, researches evaluating the effect of comprehensible input and comprehensible output on speech complexity are very limited and this was, in fact, one of the motives that encouraged the researcher to conduct this study.

Statement of the Problem

The major concern of this study is that despite the great amount of input Iranian EFL learners receive during six years of formal schooling in high school and four years of academic education in university, still a great number of Iranian University EFL learners are not able to speak English with an acceptable degree of accuracy and complexity. In fact, the problem arises from the point that in Iran, as a foreign language context, both Iranian language teachers and learners usually are more interested and proficient in receptive skills (reading and listening) than productive skills of writing and speaking. This tendency naturally

seems to result in a gap between comprehensible input (CI) and comprehensible output (CO) and a comparatively low efficiency in language production.

Objectives of the Study

Regarding the stated problem, the present study mainly aims to investigate and compare the impact and the degree of influence of comprehensible input and output on the accuracy and complexity of Iranian EFL learners' oral speech. It has also been set to determine if there exist trade-off effects between the accuracy and complexity measures of participants' oral speech.

Research Questions

The following research questions have been addressed in this study:

1- Do CI and CO affect the accuracy of English foreign language learners' speech?
2- Do CI and CO affect the complexity of English foreign language learners' speech?
3- Is CI or CO more rigorous and influential for developing the speech accuracy of Iranian English foreign language learners?
4- Is CI or CO more rigorous and influential for developing the speech complexity of Iranian English foreign language learners?
5- Does any trade-off effect exist between accuracy and complexity dimensions of English foreign language learners' oral speech?

Research Hypotheses

Based on the questions cited above, this study aims at testing the following null hypotheses:

H01: Neither CI nor CO affect the accuracy of EFL learners' oral speech.

H02: Neither CI nor CO affect the complexity of EFL learners' oral speech.

H03: Neither CI nor CO play a more significant role than the other one in the production of more accurate oral speech.

H04: Neither CI nor CO play a more significant role than the other one in the production of more complex oral speech.

H05: There exists no trade-off effect between accuracy and complexity dimensions of oral speech.

Method

Participants

This study was conducted with 60 female English foreign language students all majoring in English language teaching at MA level in Islamic Azad University, Najafabad Branch. They were between 24 and26 years old and all Persian native speakers. They were selected from a whole population pool of 80 using the listening and speaking modules of the standard test of International English Language Testing Service. Later, these homogenized participants were randomly assigned to three groups labeled as input, output, and control groups. The participants went through a pre- task phase before being engaged in the main-task phase.

Instruments

- IELTS Test

International English Language Testing Service, a standard test for speakers of other languages, was used as a reliable test for the selection (homogenization) of 60 participants with intermediate speaking ability. Since the main focus of the research was oral speech assessment, the listening and speaking modules of the test were used to assess the speakers, and reading and writing modules were deleted.

- Reading Comprehension Text

In the pre-task phase, those 20 participants who had been selected as the input group were given a reading text selected from the book "Intermediate Select Reading" that has specially been written for students at the intermediate level of English proficiency. To help students successfully tackle and understand the passage, the writer had provided the following support tools: Vocabulary glosses for challenging words and expressions, explanations for language and cultural references that appeared in blue type in the reading, and numbered lines for easy reference. The main objective of this text was to provide the participants with comprehensible input about the topic (culture shock) in the pre-task phase and later investigate its effect on the subsequent oral speech of the participants in the main-task phase.

- Standard Coding Scheme

To score the recorded monologues of participants in the main –task phase, a standard coding system was utilized. The coding process was done in terms of grammatical accuracy, and syntactic complexity of structures based on Bygate's (2001) standard coding system in which complexity is measured in terms of number of words per t-unit, where t-unit is defined as "a finite clause together with any subordinate clauses dependent on it" (Bygate, 2001, p. 35) and accuracy was measured by calculating the incidence of errors per t-unit; the higher the number, the less accurate the language (Bygate, 2001).

Procedures

To select 60 homogeneous intermediate EFL speakers for the study, 80 Persian, female EFL learners all majoring in English language teaching at MA level participated in the listening and speaking modules of the IELTS test and those whose scores were approximately 5 were selected. As the next step, these 60 participants were randomly divided into three groups of 20: The input experimental group, the output experimental group, and the placebo control group. This study consisted of two phases: The pre-task phase and the main-task phase. The members of these groups were expected to participate in activities within the pre and main-task phases.

- The Pre-Task Phase

1. The Input Group

The participants in this group in pre-task phase were asked to read a comprehension text carefully with the purpose of comprehending and getting information. The main objective of this activity was to provide the participants with CI and later investigate its effect on the subsequent oral speech of the participants. The topic of the reading text (culture shock) was exactly the same as the topic of discussion selected for the output group in the same pre-task phase in order to make sure that under identical conditions (the same age, gender, level of speaking skill, L1, the time allowed, and topic) any difference in the future oral performance of participants would be due to the mode and way through which the topic has been handled (i.e., through CI or CO).

2. The Output Group

The participants in the output group produced comprehensible output through discussing and negotiating the topic of culture shock, trying to produce the language that was both comprehensible to their partners and at the same time well formed. In the case of output group, the participants were divided into three groups of 5. Five minutes before starting the discussion, the participants were given the topic and its related sub-topics accompanied by the following instructions in written form in order to emphasize how they were expected to participate in the discussion so that the aim of this phase that was providing comprehensible output would be achieved:

1. Look at the presented topic and sub-topics and think about them for five minutes.
2. The participation of every individual in the discussion is important and necessary.
3. Try to understand what others say, using any technique or strategy you have at your disposal.
4. Try to speak in such a way that others will understand you, and if any misunderstanding occurs, try to compensate it;

for example, by rephrasing, paraphrasing, simplifying your speech, etc.

5. Mutual understanding and peer correction is also quite welcomed.

The Control Group

In this phase, the members of the control group were not engaged in any particular activity and merely attended their regular classes.

- The Main-Task Phase

In the main-task phase, the participants from all three groups were asked to perform a loud monologue within 10 minutes about the topic that was thematically related to the topic of the pre-task phase. The main purpose of this activity was to determine whether the group that has received comprehensible input about the topic (the input group) could produce more accurate and complex oral speech in this phase or the group that had practically put into practice their inter-language knowledge through discussing and negotiating the topic (the output group).

The oral performance of participants of each group was recorded on three separate CDs. Later, the researcher and two other raters listened to the recordings and transcribed them carefully. Keeping in mind the main objective of this study that was investigating the effect of CI and CO on the grammatical accuracy and syntactic complexity of oral speech, these transcriptions had to be coded in terms of accuracy and complexity. This was done using Bygate's (2001) standard coding system, where accuracy was accounted by the number of errors per t-unit and complexity by the number of words per t-unit. To ensure inter-coder reliability, the coding process was done by the researcher herself and two other competent raters. Correlation of the obtained scores was found to be 0.89, confirming the objectivity of the ratings. Rating the coded transcriptions was the last step. Those who had made fewer grammatical errors in each t-unit, received higher scores in accuracy and those who had made longer sentences and structures (i.e., used more words per t-unit) received higher scores in complexity. Then, these scores were submitted to statistical analysis.

Data Analysis

In the present study, one-way ANOVA and Post Hoc tests were performed to examine if there exist any significant differences among three input, output, and control groups regarding the accuracy and complexity of speech and if such difference exists, is the observed critical difference more meaningful in the case of input or output group. In addition, correlation coefficient was run between the accuracy and complexity scores within each group in order to investigate the relation between these two measures of oral speech and to find the existence or non-existence of trade-off effects.

Accuracy of Speech and the Role of CI and CO

Our research question was whether comprehensible input and output have any effect on the accuracy of EFL learners' oral speech. The purpose of this question is to find out whether these two variables result in more accurate speech or not. To address this question, the number of grammatical errors per t-unit was considered, and the generated errors of participants in three groups were computed using ANOVA. Table 1 presents the descriptive statistics of each group.

Table1
Descriptive Statistics: Speech Accuracy among Groups

Descriptives

accuracy of speech

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
input	15	15.07	6.158	1.590	11.66	18.48	5	25
output	15	9.80	4.411	1.139	7.36	12.24	4	18
control	15	17.33	6.694	1.728	13.63	21.04	5	28
Total	45	14.07	6.535	.974	12.10	16.03	4	28

As the descriptive data in Table 1 show, the accuracy mean scores of participants both in the input and output group (**15.07** and **9.80**) are lower than those of the control group (**17.33**). As it has been mentioned earlier, in the case of accuracy measurement that is actually an inaccuracy measurement in this study, the smaller the figure, the better the results. So, the lower mean scores of the input and output group in comparison with the control group indicate that the participants in these two groups have made fewer errors in their speech.

According to the obtained means, the accuracy mean scores in the input and output groups are different from and lower than the control group, indicating that CI and CO have helped to reduce the incidence of errors and hence the accuracy of learners speech: the participants have had the fewest number of errors in the output group, followed by the input group, and then the control group. The results of ANOVA, shown in Table 2 largely reflect this trend:

Table 2

ANOVA: The Accuracy of Speech

ANOVA

accuracy of speech

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	448.133	2	224.067	6.578	.003
Within Groups	1430.667	42	34.063		
Total	1878.800	44			

The findings reveal a significant difference among the participants' accuracy measures in the three groups. (F=6.578, df=2, P=.05), meaning that there were significantly different numbers of errors in each group and that both CI and CO have affected the accuracy of EFL learners' speech. In other words, this significant difference shows that CI and CO have been two influential factors reducing the number of errors as indicators of accuracy in this research. Based on the results presented in Tables

1 and 2, the first research hypothesis denying the effect of CI and CO on speech accuracy is rejected.

The Complexity of Speech and the Role of CI and CO

Similarly, the second research question investigated the effect of CI and CO on the complexity of EFL learners' oral speech. To address this question, the participants speech complexity measured through the number of content words per t-unit were compared among three groups using ANOVA and post hoc tests. Table 3 presents the descriptive statistics for the number of words per t-unit indicating speech complexity of the participants in each group:

Table 3

Descriptive Statistics: Speech Complexity among Groups

Descriptives

complexity of speech

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
input	15	496.60	67.144	17.336	459.42	533.78	392	621
output	15	462.53	65.157	16.823	426.45	498.62	388	578
control	15	389.67	69.813	18.026	351.01	428.33	311	524
Total	45	449.60	79.816	11.898	425.62	473.58	311	621

Complexity mean scores of the participants in the input and output groups (**496.60 and 462.53**) are different from and higher than that of the control group (**389.67**), indicating that CI and CO have affected speech complexity. In fact, the existence of difference among the complexity mean scores of the participants in the three groups implies that input reception and output production tasks have led to changes in the number of words used per t-unit ; hence the complexity of speech.

As the results show, the participants in the input group have produced more words than either the participants in the output or control group. The result of ANOVA presented in Table 4 confirms this finding. This test was run to determine if the difference among the three involved groups regarding the complexity of speech was significant or not.

Table 4
ANOVA: Complexity of Speech

ANOVA

complexity of speech

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	89524.133	2	44762.067	9.854	.000
Within Groups	190784.667	42	4542.492		
Total	280308.800	44			

As the results of Table 4 show, there existed a significant difference among the participants' complexity measures in three groups. **(F= 9.854, df=2, p=.05)**. It means that CI and CO in comparison with the control group, have affected the participants' complexity measure that the degree and significance of each of these variables on developing speech complexity will be later compared and determined through post hoc test.

Comparing the Effect of CI and CO on Developing Speech Accuracy

In order to investigate the third research question pertaining to the comparison between the effects of CI and CO on speech accuracy, the ANOVA analysis of accuracy presented in Table 2 was followed by post hoc Scheffe test. This post hoc test was run to determine whether CI or CO plays a more significant role in developing speech accuracy. The results of the post hoc test have been presented in Table 5:

Table 5
Post Hoc Test: Analysis of Accuracy

Multiple Comparisons

Dependent Variable: accuracy of speech

Scheffe

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
input	output	5.267	2.131	.058	-.14	10.67
	control	-2.267	2.131	.572	-7.67	3.14
output	input	-5.267	2.131	.058	-10.67	.14
	control	-7.533*	2.131	.004	-12.94	-2.13
control	input	2.267	2.131	.572	-3.14	7.67
	output	7.533*	2.131	.004	2.13	12.94

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

As shown in Table 5, the results revealed that the difference existing between the numbers of errors (as indicator of accuracy) in the CI and CO groups was not significant. This means that the influence of these two variables on developing speech accuracy was not significantly different and neither was more rigorous than the other one in decreasing the number of errors. Thus, the third research hypothesis which stated that there is no difference between the effects of CI and CO on the accuracy of Iranian EFL learners' oral speech is confirmed.

Comparing the Effect of CI and CO on Developing Speech Complexity

Research question four asks whether there exist differences between the effects of CI and CO on the speech complexity of EFL learners. To address this question, the ANOVA test that was run for complexity measure in Table 4 was followed by the post hoc Scheffe test to determine if CI or CO plays a more influential role in improving speech complexity. Table.6 shows the results of this post hoc test

Table 6
Post Hoc Test: Analysis of Complexity

Multiple Comparisons

Dependent Variable: complexity of speech

Scheffe

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
input	output	34.067	24.610	.392	-28.39	96.52
	control	106.933*	24.610	.000	44.48	169.39
output	input	-34.067	24.610	.392	-96.52	28.39
	control	72.867*	24.610	.019	10.41	135.32
control	input	-106.933*	24.610	.000	-169.39	-44.48
	output	-72.867*	24.610	.019	-135.32	-10.41

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

This current analysis directly compares speech complexity between groups two by two in an effort to explore whether CI or CO is more influential in developing complexity. The results revealed that the difference existing between number of words (as indicator of complexity) in the CI and CO groups was not significant, that is neither of these variables proved to be more rigorous and influential than the other one in developing the complexity of Iranian EFL learners' oral speech. Therefore, the fourth research hypothesis that stated no difference exists between the effects of CI and CO on the complexity of oral speech is confirmed.

The Existence of Trade-off Effects between Accuracy and Complexity

Complexity and accuracy, as two measures of oral speech, were compared between and among input, output, and control groups using ANOVA and post hoc Scheffe test. In order to investigate the relation between these two measures within each group, the Pearson correlation was run between the error scores (as

indicator of accuracy) and complexity scores. The following table shows the results.

As the Table 7 shows, there exists a significant correlation between the set of error scores (number of errors per t-unit) and complexity scores (number of words per t-unit) within the input group.

Table 7
Correlation between Accuracy and Complexity Scores within the Input Group

Correlations

		accuracy of speech	complexity of speech
accuracy of speech	Pearson Correlation	1	.959**
	Sig. (2-tailed)		.000
	N	15	15
complexity of speech	Pearson Correlation	.959**	1
	Sig. (2-tailed)	.000	
	N	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

In order to determine if this correlation is positive or negative, we present the following figure:

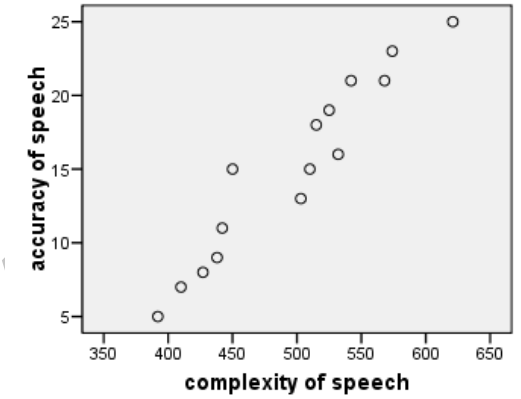

Figure 1. Correlation between Accuracy and Complexity Scores within the Input Group

As shown in Figure 1, from the upward movement of the graph it is understood that the correlation is positive and

consequently increase in the number of errors is accompanied by increase in the number of words and vice versa. However, as it has been mentioned earlier, in the case of accuracy measurement that is actually an inaccuracy measurement in this study, the lower the error score, the higher the accuracy and the higher the error score, the lower the accuracy. In fact, this positive correlation exists between the two sets of scores one representing number of errors and the other one the number of words. But considering the diverse relation between the actual accuracy and the error score (as indicator of accuracy), we draw the conclusion that the positive correlation shown in the figure 1 is between complexity and error scores, that is to say, the correlation between actual accuracy and complexity of oral speech is negative.

To investigate the relation between accuracy and complexity scores within the output group, again the Pearson correlation is used. Table 8 presents the obtained results:

Table 8

Correlation between Accuracy and Complexity Scores within the Output Group

Correlations

		accuracy of speech	complexity of speech
accuracy of speech	Pearson Correlation	1	.974**
	Sig. (2-tailed)		.000
	N	15	15
complexity of speech	Pearson Correlation	.974**	1
	Sig. (2-tailed)	.000	
	N	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

As it can be observed in Table 8, there exists a significant correlation between the number of errors and number of words made by individuals within the output group. However, whether this correlation is positive or negative is not apparent from the

table itself, and for this purpose, the obtained results need to be presented through Figure 2.

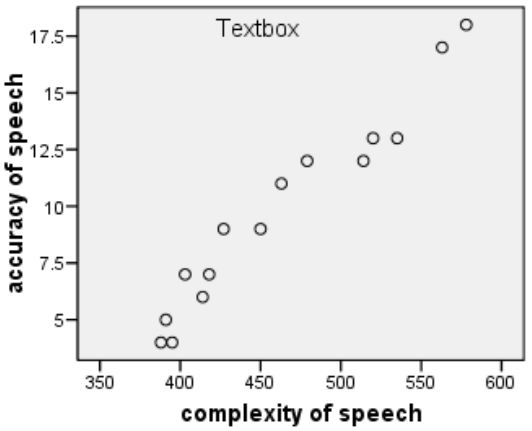

Figure 2 Correlation between Accuracy and Complexity Scores within the Output Group

The upward direction of graph indicates that the correlation between the error and complexity scores within the output group is positive and consequently the increase in the number of errors is accompanied by the increase in complexity and vice versa. However, since in this study accuracy is an inaccuracy measurement and the relation between the actual accuracy and error score, as indicator of accuracy, is diverse, it can be concluded that the correlation between complexity and actual accuracy is a negative one. Pearson Correlation is again used to determine if there exists any significant correlation between accuracy and complexity scores within the control group:

Table 9 shows that like the input and output group, the correlation between these two sets of scores in the control group is also significant.

Table 9

Correlation between Accuracy and Complexity Scores within the Control Group

Correlations

		accuracy of speech	complexity of speech
accuracy of speech	Pearson Correlation	1	.882**
	Sig. (2-tailed)		.000
	N	15	15
complexity of speech	Pearson Correlation	.882**	1
	Sig. (2-tailed)	.000	
	N	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

Drawing Figure 3 can help to determine if such significant correlation is positive or negative.

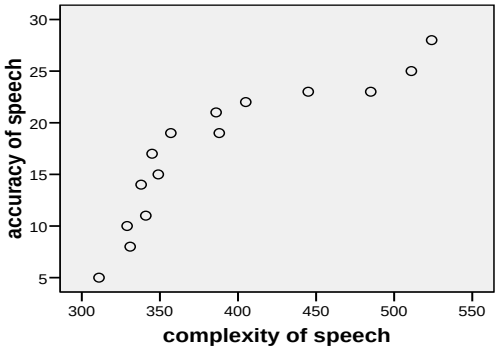

Figure 3 Correlation between Accuracy and Complexity Scores within the Control Group

Upward direction always signifies a positive correlation and downward direction a negative one. Based on this Figure, in the control group, the correlation between error and complexity scores is positive. But, as mentioned before, due to the converse relation between error score (as indicator of accuracy) and the actual accuracy, it is concluded that there exists a negative correlation between the actual accuracy and complexity as two measures of oral speech.

Discussion

The first null hypothesis of this research aimed at determining if comprehensible input and output have any effect on the accuracy of EFL learners' oral speech. The findings of the present study showed that both input reception and output production tasks caused changes in the syntactic accuracy of EFL learners' speech. This finding of the study supports previous claims made concerning the influential role of CI and CO in developing the accuracy measure of oral performance.

As stated by Krashen (1985), speaking skills and the knowledge of grammar rules occur as long as sufficient amount and type of input is provided. In fact, in approaches supporting CI, speech accuracy is thought to evolve as the learner progresses and receives more and more CI, without conscious attention to form or error correction. On the other hand, great numbers of experiments have also proved the influential role of CO in developing speech accuracy and review of the literature provides us with convincing theories as why comprehensible output can improve the speech accuracy of EFL learners. Three specific functions have been mentioned for comprehensible output that is believed to be different from and independent of comprehensible input. 'Noticing function', 'hypothesis testing function' and 'meta-cognitive function' are regarded as three significant functions of CO that can help learners move from semantic to syntactic processing and as the result develop their grammatical accuracy. Studies done about comprehensible output and each of its specific functions positively support and confirm the role it plays in improving speech accuracy of EFL learners. There exist great number of studies in the literature that have provided support for the claim that pushing learners to produce comprehensible output results in gains in speech accuracy (e.g., Swain, 1985, 1995; Nagata, 1998; Tarone & Liu, 1995; Izumi, 2002; Dekeyser & Sokalski, 1996; Mackey, 1999; Noubishi & Ellis, 1993). Results of these studies support the link between output production and in particular the development of grammatical accuracy. Therefore, based on the created change and development in speech accuracy of input and output group

participants, regardless of the degree of significance of such change in each group, the first null hypothesis stating that CI and CO have no effect on speech accuracy is rejected.

The second null hypothesis of the research similarly focuses on determining whether CI and CO play any role in developing speech complexity. The review of the literature revealed that the number of studies which directly investigated the effect of comprehensible input and output on the complexity of oral speech is very limited and fewer than those investigating the impact of these factors on speech accuracy. As the results of this study revealed, comprehensible input and output both created changes in speech complexity of participants. Determining whether CI or CO was more rigorous, influential, and significant in creating such a change and development in speech complexity is the matter that is considered when evaluating the fourth hypothesis of the study. However, the created change in speech complexity of input and output group members, regardless of the degree of significance of such change in each group, caused the second null hypothesis denying the effect of comprehensible input and output on developing speech complexity to be rejected as well.

The third null hypothesis of the research states that comprehensible input and output have the same impact on the accuracy of EFL learners' oral speech. The results obtained from this study confirmed this null hypothesis. However, the review of the existing literature reveals that previous studies conducted in this regard are not in line with this result of the study and previous experiments mainly confirm a more influential role for comprehensible output in developing speech accuracy and formal linguistic features and a less influential role for comprehensible input in this regard.

Beside others, Nobuyoshi and Ellis (1993), for example, claim to have provided data showing that comprehensible output results in actual improvement in speech accuracy. They stated that their study " provides some support for the claim that pushing learners to improve the accuracy of their production results in not

only immediate improved performance but also in gains in accuracy over time" (p. 208).

On the other hand, one of the most significant experiments that proves the lack of success of CI in improving the accuracy of speech is the case of learners in the immersion context, where the learners were exposed to abundant, extensive amount of comprehensible input in both written and spoken form, but still produced oral speech that was full of grammatical errors. Ryan (n.d.) refers to the study conducted with Canadian immersion students, in which Swain (1995) has shown that even though students had received abundant comprehensible input in French and were somewhat fluent in the language, they still had not acquired grammatical competence in the language. Immersions students' many syntactical errors in French confirmed that the target language grammatical system had not been fully acquired.

One underlying reason why in this research neither of these variables succeeded to be more influential in improving speech accuracy may be the way of conducting tasks and mainly the way comprehensible output was put into practice. There is the probability that during the output production task, the CO in its real sense that calls for negative feedback, modification, restatement, and correction had not been truly practiced and consequently the main function of output that is moving learners from semantic to syntactic processing had not been applied. In fact as Krashen (1994) has stated, output in this sense is scarce, so it is possible that in this study, too, the real output that was expected to be produced had not been practiced.

The fourth null hypothesis of the research that states comprehensible input and output have the same effect on the complexity of EFL learners oral performance is also confirmed based on the results obtained from this study. As the obtained results show, complexity level was not significantly different in the speech of input and output group learners.

As it was mentioned before, the numbers of studies that have investigated the effect of CI and CO on the development of speech complexity are very limited and consequently there do not exist enough experiments in this regard to be compared with the results

of this study. However, in seeking for an underlying reason, we can again mention the way in which CO was put into practice had limitations and the real sense of CO, i.e. pushing learners from a semantic to syntactic processing has not been achieved.

The fifth null hypothesis of this study investigated the existence of trade- off effects between accuracy and complexity within all three groups of the study. Analyzing the scores of individual participants revealed that those participants who performed oral speech with higher accuracy showed less gains in complexity, and those who had higher scores in speech complexity, were weaker in performing accurate, grammatical speech. Such results are quite in line with previous experiments which were concerned with accuracy, complexity, and fluency as three measures of EFL learners' speech production. According to Skehan and Foster (2001), attentional resources are limited and attending to one aspect of performance may mean that other dimensions are neglected. They propose that for language development to proceed optimally, a balance needs to be established between these three performance dimensions. Mentioning some of these studies at this point can be a good evidence for proving such claim.

Foster and Skehan (1996) examined the influence of task type and degree of planning on three different aspects of L2 performance: fluency, accuracy, and complexity. The study employed three types of tasks (personal information exchange, narrative, and decision-making) under three planning conditions (unplanned, planned but without detail, and planned with detail). In their discussion, Foster and Skehan noted that a trade-off existed between the goals of performance complexity and performance accuracy. They explained that individuals have a limited capacity for attention, as noted earlier, so when a task is more cognitively demanding, attention is diverted from formal linguistic features-the basis of accuracy-to dealing with these cognitive requirements.

In Brazil, D'Ely (2006) investigated the impact of four types of pre-task planning on the L2 oral performance of 47 learners of English. Participants' speaking was assessed in terms of fluency, accuracy, complexity, and lexical density. Results of the statistical

analyses conducted showed that different planning conditions affected learners' performance in different ways. D'Ely's results also provided evidence for the trade-off among fluency, accuracy, complexity, and lexical density due to limitations in attentional resources during speaking, giving support to Foster and Skehan (1996).

Conclusion

Based on the findings of this study, comprehensible input and output are both influential variables for improving the accuracy and complexity measures of L2 learners' speech. Significant changes occurred in the performance of input and output group participants in terms of accuracy and complexity after being engaged in input reception and output production tasks, indicating that comprehensible input and output are two variables that can improve speech accuracy and complexity.

The results of this study also revealed that there existed no significant difference between the number of errors (as indicator of accuracy) in the speech of input and output group members, indicating that neither CI nor CO has been more influential in developing the accuracy dimension of Iranian EFL learners' oral speech. To put it differently, the decrease in the number of errors was not so much different in the speech of input and output group participants. Similarly, the lack of noticeable difference between input and output group members regarding the increase in the number of words (as indicator of complexity) in their speech provided evidence that neither CI nor CO has been more effective in improving the speech complexity of Iranian EFL learners. This result of the study, however, is not in harmony with the previous experiments as the majority of studies conducted in this regard mainly have reported a more effective role for comprehensible output in developing accuracy dimension of oral speech in comparison with comprehensible input.

Another obtained result of this study that can be matched up with recent theories is the trade-off effects between accuracy and complexity measures of oral speech. Such trade-off effect means that one of these variables progresses at the expense of the other

one. In the present study, within all three groups, when attention was devoted to speech complexity, it was diverted from formal linguistic features (accuracy), and, on the other hand, accuracy progressed at the expense of speech complexity.

Implications of the Study

One possible pedagogical implication that can be derived from the findings of this study is that in the course of L2 speech development, teachers should design and manipulate tasks in such a way that enough practice is allowed in various dimensions of oral speech. In other words, since different tasks provide different opportunities for practicing specific aspects of L2 speech production, various tasks need to be manipulated to enable learners practice various dimensions of speech production, focusing both on meaning and on form and in focus on form, according to Finardi (2008, P.9), "a balance should be made between hypothesis testing and restructuring (complexity) and the control of stable elements in the interlanguage system (accuracy)". Generally it is implied that the most effective way for improving oral speech, based on the literature and the results obtained from this study, is an eclectic approach which conflates both CI and CO.

To use Bygate's (1999) words: "Feed people with narrative tasks and they will crunch up some aspects of speech in one way, sharpening certain linguistic teeth, i.e., cognitively mapping certain types of language and certain types of communicative demand. Feed them different tasks, and different linguistic teeth might be developed."(p. 39).

Limitations of the Study

The most notable limitation is related to the IELTS test used to select homogeneous intermediate learners. This test was selected from an IELTS course book, i.e., Cambridge IELTS 5, due to the problem of accessibility to the real IELTS test. So there is the possibility that the obtained results have not truly reflected the proficiency level of the participants. Moreover, based on this course book test, the score of participants was on average 5, but

more specifically 4.5 to 6. Thus, the oral performance of learners was likely to have been influenced by this broad range of scores.

Due to the problem of accessibility, like most SLA studies, this study also involved a small number of participants. The present research was conducted with 45 participants, 15 members in each group. To be able to generalize these results more confidently, a larger number of participants are required.

One aspect that must be taken into consideration when analyzing the results of this study is the fact that since the participants who cooperated in this research came from an experimental group, there is the probability that some of them may have perceived the speaking task as tests and therefore behaved accordingly. As Iwashita, McNamara, and Elder (2002) suggest, performance on tests differ from performance in class and so it has to be analyzed differently and with caution. As stated by Finardi (2008), task implementation for research purpose must be carried out with care and consideration of these issues.

Suggestions for Further Research

The current study population involved intermediate-level learners of English as a foreign language with Persian as their L1. Conducting the study with another different population regarding their L1, level of proficiency, educational background, learning environment, and other variables may lead to quite different results about the influence of input and output on accuracy and complexity of EFL learners' oral speech.

In the present study, the output hypothesis was put into practice through interactions between L2 learners, rather than among L2 learners and native speakers. Thus, the results of the current study should not be generalized to native speaker-learner interactions without additional research.

The main focus of the present study is on the accuracy and complexity of oral speech. However, the effect of CI and CO can also be investigated regarding the fluency of speech.

The mode of presenting CI to learners (either through listening or reading) may influence their comprehension and

consequently their following performance. So research on the role and effect of mode of input presentation seems necessary.

The Authors

Omid Tabatabaei is an assistant professor of Islamic Azad University, Najafabad Branch, English Department. He is currently the Vice-Dean of Humanities Faculty and the head of English Department as well. He has published and presented articles in national and international journals and conferences. His areas of interest are: Task-based instruction, L2 acquisition, assessment and testing, and psycholinguistics.

Masume Yakhabi has received her MA in TEFL from Islamic Azad University, Njafabad Branch. She is teaching General English and ESP courses at Azad and Payame Noor Universities. Her major research interests are L2 acquisition and teaching methodology.

References

Bygate, M. (1999). Task as context for the framing, reframing, and unframing of language. *System*, *27*, 33-48.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning and testing* (pp. 37-53). London: Longman.

Dekeyser, R, & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning*, *46*(4), 613-642.

D'Ely, R. C. F. (2006). A focus on learners' meta-cognitive processes: The impact of strategic planning, repetition, strategic planning plus repetition, and strategic planning for repetition on L2 oral performance. Unpublished Doctoral Dissertation, University of Santa Catarina.

Ellis, R. (2008). Principles of instructed second language acquisition. CALL(Center for Applied Linguistics). Retrieved September 25, 2009 from http:// www.cal.org/resources/digest/instructed2ndlang.html.

Faerch, C., & Kasper, G. (1986). The role of comprehension in second language learning. . *Applied Linguistics*, 7(3), 257-274.

Finardi, R. K. (2008). Effects of task repetition on L2 oral performance. *Scielo, 47(1).* Retrieved November 25, 2009 from http:// www.scielo.br/scielo.php?script=sci-arttext&pid=5do3-18132008000100003&hg=es&nrm=iso&tlng=en

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*, 229-323.

Fulcher, G. (2003). Testing second language speaking. TESL, 8(1). Retrieved June, 2004 from http:// tesl-ej.org/ej 29/r5.html.

Harati, N. (2000). The impact of recast versus negotiation of form on the grammatical accuracy of Iranian EFL students' speech. A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Teaching English as a Foreign Language. In the Foreign Languages Department of Iran, University of Science and Technology.

Iwashita, N. & McNamara, T. & Elder, C. (2002). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language learning*, *51*(3), 401-436.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition, 24*, 541-577.

Krashen, S. D. (1981). *Principles and practice in second language acquisition: English Language Teaching Series*. London: Prentice-Hall International (UK) Ltd.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London: Academic Press.

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*, 557-587.

Mangubhai, F. (2001). Book floods and comprehensible input floods: Providing ideal conditions for second language acquisition. *International Journal of Educational Research, 35,*147-156.

Nagata, N. (1998). Input vs. output practice in educational software for second language acquisition. *Language Learning & Technology, 1*(2), 23-40.

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

Ryan, J. (n.d.). A review of the role of the output in second language acquisition with anecdotal examples from a Japanese learner's experience. *Language Learning, 38*(17), 42-54.

Schutz, R. (2007). Stephen Krashen's theory of second language acquisition.Retrieved November, 2007 from http://www.sk.com.br/sk-krash.html.

Shehadeh, A. (2002). Comprehensible output, from occurrence to acquisition: an agenda for acquisitional research. *Language Learning, 52*(3), 597-647.

Skehan, P . (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. & Foster, P. (1997). Task type and task processing conditions. *Language Teaching Research, 3*(1), 185-211.

Skehan, P. & Foster, P. (2001). Cognition and Tasks. In Robinson, P. (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University Press.

Soleimani, H. & Ketabi, S. & Talebinejad, M. (2008). The noticing function of output in acquisition of rhetorical structure of contrast paragraphs of Iranian EFL university students. *Linguistic Online.* Retrieved April, 2008 from http:// www. linguistic-online.de/34-08/soleimani.html

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In Gass,S; & Madden, C. (EDs), *Input in second language acquisition* (pp.235-256). Newyork. Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seildlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of H.G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2008). The output hypothesis: its history and its future. A keynote speech in The 5th Annual International Conference on ELT in Beijing. Retrieved March 1, 2008 from http://www.celea.org.cn/2008/keynote/ppt/merrill%20Swain.pdf.

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.

Tarone, E., & Liu, G.Q. (1995). Situational context, variation, and second language acquisition theory. In G. Cook & B. Seidelhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp.107-124). Oxford: Oxford University Press.