TELL, Vol. 8, No. 1, Spring & Summer 2014, pp. 141-166

Short-term and long-term impacts of individual and collaborative pragmatic output on speech act production

Zia Tajeddin
Professor of Applied Linguistics, Allameh Tabataba'i University

Marzieh Bagherkazemi[1]
Ph.D. Candidate, Allameh Tabataba'i University

Received October 1, 2013
Accepted June 5, 2014

Abstract

Since the 1990s, a substantial body of interventionist interlanguage pragmatics (ILP) research has probed the efficacy of various instructional pragmatics approaches, with speech acts serving as its prime target. The present study compared the short-term and long-term effects of individual and collaborative output-based instruction on EFL learners' acquisition of English "apologies," "requests," and "refusals." For this purpose, 54 intermediate EFL learners, making up an individual output group (N=26) and a collaborative output group (N=28), participated in the study. Individual and collaborative output-based instruction was offered over nine consecutive sessions (three sessions on each speech act). Each of the

[1] Corresponding author at: Allameh Tabataba'i University
E-mail Address: mbagherkazemi@gmail.com

nine treatment sessions involved the presentation of video input containing the speech acts under investigation, followed by individual and collaborative output production and manipulation tasks. Moreover, a 24-item WDCT was used to measure the participants' speech act production ability at the pre-treatment, immediate post-treatment, and delayed post-treatment phases of the study. The results indicated significant gains for both individual and collaborative output groups from the pretest to the immediate and delayed posttests, and the greater efficacy of collaborative output over individual output. The findings reveal the potential of learner output, both individual and collaborative, for ILP development and the greater advantage collaborative output production and manipulation tasks can offer.

Keywords: collaborative output, comprehensible output hypothesis, individual output, interlanguage pragmatics, speech acts

1. Introduction

Within the broad domain of second language acquisition, "pragmatic competence" explicitly premiered in Bachman's (1990) model of communicative competence, underscoring the relationship between "language users and the context of communication" (p. 89). Since the 1990s, a great many studies have compared the effects of various instructional approaches on the acquisition of L2 pragmatic competence (e.g., Bardovi-Harlig, 2001; Kubota, 1995; Li, 2012; Rose & Ng Kwai-fun, 2001; Tajeddin, Keshavarz, & Zand Moghadam, 2012; Takimoto, 2007). Although the existing research has been almost exclusively preoccupied with explicit vs. implicit instruction, findings are yet inconclusive. In their meta-analyses of instructional pragmatics, Jeon and Kaya (2006) and Takahashi (2010) showed that explicit instruction has generally proved to be more effective than implicit instruction. However, they regret the

dearth of research on instructional pragmatics, which precludes the achievement of any conclusive evidence.

Equivocal research results might be due to the fact that explicit and implicit instruction can be realized in a variety of ways, and it might not be the implicitness or explicitness of the approach that matters the most. Given this, it seems to be insightful to unleash instructional pragmatics research from this predominant binary distinction, and investigate what learners themselves can bring to the task of acquiring pragmatic competence. One potential framework within which pragmatic competence can be investigated from an acquisitional learner-oriented perspective is Swain's (1985) *comprehensible output hypothesis*, which capitalizes on the significance of providing learners with opportunities for classroom language use and pushing them to modify their output and make themselves more comprehensible (Mitchell & Myles, 2004; Shehadeh, 2002). Since Swain (1985) put forth her hypothesis, a large number of studies have been conducted to substantiate claims as to the significance of L2 learners' output for their interlanguage development (e.g., Adams, Nuevo, & Egi, 2011; Egi, 2010; Ellis & He, 1999; Izumi, 2002; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Nassaji & Tian, 2010; Storch, 2005, 2007). However, they have mainly addressed grammar and vocabulary acquisition, and the role of learners' individual and collaborative output in developing their L2 pragmatic competence is yet to be demonstrated. The present study was designed to address this research gap, i.e. impacts of individual output (IO) and collaborative output (CO) on the production of speech acts.

2. Literature Review

There is ample ILP research evidence as to the teachability of pragmatics and the effectiveness of pragmatic instruction vis-à-vis simple exposure (see Taguchi, 2011 for a review). As a consequence, prominently featuring in ILP research are comparative studies of various instructional pragmatics approaches targeting speech acts, including inductive or deductive explicit instruction (e.g., Rose & Ng Kwai-fun, 2001), input enhancement (e.g.,

Takahashi, 2001), task-based instruction (e.g., Tajeddin, Keshavarz, & Zand Moghadam, 2012), processing instruction (e.g., Takimoto, 2007), practice-based instruction (e.g., Li, 2012), and output-based instruction (e.g., Jernigan, 2007). Studies of output-based instructional pragmatics are far and few between, and no ILP study has ever compared effects of individual and collaborative output on L2 learners' speech act production ability. This section offers an overview of the comprehensible output hypothesis and research on individual and collaborative output-based instruction.

In the face of the dominant information-processing SLA paradigm of the 1980s, the "comprehensible output hypothesis" was advanced by Swain (1985), to counter Krashen's (1985) "input hypothesis." Central to Swain's hypothesis is the concept of "pushed or modified output," denoting the idea that the value of learner output is not mainly in its functioning as fluency practice (Swain, 1995). Rather, when producing the target language, L2 learners are *pushed* to make their output more accurate, appropriate, and comprehensible, especially when they have difficulty getting their message across. It is this potential that contributes to the restructuring of L2 learners' interlanguages (Shehadeh, 2002). Moreover, Swain (1995) refers to "noticing," "hypothesis-testing," and "reflection" as the three prime functions of learner output, lending it an acquisitionally significant role. Later, Swain (2000, 2006) (1) replaced "output" with the more telling label of "languaging," stating that it better captures the idea of language use as a cognition mediating tool, and (2) introduced the concept of "collaborative dialog" which she defined as follows:

> It is dialogue that constructs linguistic knowledge. It is what allows performance to outstrip competence. It is where language use and language learning can co-occur. It is language use mediating language learning. It is cognitive activity and it is social activity. (Swain, 2000, p. 97)

Implied in studies of learner output is the distinction between individual and collaborative output. The thrust to push one's output, according to Swain and Lapkin (1995), is created by either *internal* or *external* feedback which contextualized language use leads to. In

other words, modified output can be either individual/self-initiated, which is equivalent to "monitoring," or collaborative/other-initiated, which is on a par with "uptake" (Ellis, 2008). Acquisitional SLA research on learner output has explored the efficacy of individual or collaborative output in comparison with control and/or input-based conditions (e.g., Adams, Nuevo, & Egi, 2011; Ellis & He, 1999; Izumi, 2002; Shehadeh, 1999; Smith, 2005). Such research has also explored, albeit to a less extent, individual and collaborative output in comparative studies (e.g., Nassaji & Tian, 2010; Jabbarpoor & Tajeddin, 2013; Storch, 2005, 2007), mainly targeting grammar and vocabulary. These studies have produced mixed results, with some detecting no significant difference between individual and collaborative output (e.g., Nassaji & Tian, 2010; Storch, 2005, 2007), and others substantiating the superiority of collaborative output (e.g., Jabbarpoor & Tajeddin, 2013).

Theoretically, the part played by comprehensible output in acquisitional pragmatics has not been clearly delineated. Mapping Swain's comprehensible output hypothesis onto interlanguage pragmatics, Jernigan (2007) roughs out its implications in terms of the three functions of pushed output:

1. *The noticing function:* Learners' failed attempts at producing speech acts or other pragmatic features of the target language leads them to notice their inability to appropriately convey their intended meaning, and this primes them to notice relevant features in the subsequent input or to search for more appropriate options in their own interlanguage system.
2. *The hypothesis testing function:* Through the internal or external feedback that learners receive on the pragmalinguistic or sociopragmatic aspects of their speech acts, they engage in individual or collaborative trial-and-error episodes to modify their output.
3. *The reflective function:* When pushed to the limits of their pragmatic productive ability, learners use language, either in the form of private speech or collaborative dialoging, to reflect on the pragmalinguistic and sociopragmatics aspects of their production.

In fact, ILP studies which have incorporated learner output in their design appreciate its potential only for knowledge proceduralization, and as such are implicitly or explicitly founded on the "skill acquisition" theories of language acquisition (Taguchi, 2011). These studies assume that once L2 learners are provided with information on a target language feature, ample communicative practice can assist them in their transition from controlled or declarative to automatic or proceduralized knowledge (Ellis, 2008; Li, 2012; Taguchi, 2011). However, the potential of learner output, Swain (1995) claims, goes beyond developing processing control over already acquired forms: Output can aid the acquisition or internalization of those forms in the first place. In addition, a number of empirical studies have exploited learner output as a means to investigate other relevant factors in the context of interaction. As a consequence, the efficacy of individual output has not been duly addressed. As for collaborative output, Liddicoat and Crozet's (2001) study of the effect of instruction on the acquisition of French interactional norms involved role plays, but only after a metapragmatic awareness-raising phase. Likewise, Fukuya and Zhang's (2006) study engaged learners in the collaborative production of requests, though not as the main focus of investigation. The results of the study, nonetheless, showed that collaborative output opportunities, in the absence of relevant input or feedback, would not aid learners in acquiring the full range of "request" head act strategies and their associated grammatical features.

Empirical evidence for the efficacy of collaborative output comes from Bardovi-Harlig and Salsbury (2004), who studied the development of oppositional talk in interactional exchanges between ESL learners and also between ESL learners and native speakers. The results indicated the learners' development from direct to more target-like and elaborate oppositional talk. Bardovi-Harlig and Salsbury attribute the results to the provision of conversation opportunities in the classroom which can potentially enhance ILP development even in the absence of instruction. Therefore, this study can serve as empirical support for the positive effect of "collaborative output" opportunities on ILP development.

A study which has investigated developmental pragmatics explicitly from an individual comprehensible output perspective is that of Jernigan (2007). Jernigan (2007, p. 111) investigated whether "there was an effect for an output-based instructional treatment that included pragmatic video vignettes on adult ESL learners' developing L2 pragmatic perception and production," specifically targeting "request refusals," "compliment responses," "advice giving," and "invitation." Following the presentation of a number of speech act-contained vignettes and their scripts, the +output experimental group engaged in individual written and oral text reconstruction tasks, while the -output control group answered comprehension questions. Jernigan's study revealed a significant main effect for individual pushed output opportunities on interlanguage pragmatic development, though this effect was mediated by the nature of the speech act in question, i.e. the degree of face work it involved and the modality of the output produced. The more face threatening the enactment of a speech act, the greater the amount of pressure involved in its production would be. On the other hand, written output opportunities tended to alleviate the issue to a considerable extent.

Overall, there is a critical shortage of interventionist ILP research on the impact of individual and/or collaborative output-based instruction. Moreover, the few existing studies have produced mixed findings owing to operationalizing learner output not as a main instructional element but as subordinate to input. Moreover, no ILP study has ever probed the effectiveness of individual and collaborative output in a comparative design.

3. Purpose of the Study

The present study was designed to compare the short-term and long-term effects of individual and collaborative output-based instruction on EFL learners' production of the three speech acts of "apology," "request," and "refusal." Accordingly, the following question was formulated:

> Do individual and collaborative output-based instructional approaches significantly differ with each other in terms of their

short-term and long-term effects on EFL learners' speech act production ability?

4. Method

This section provides detailed information on the participants, instructional materials, instructional treatments, data collection procedure, and data analysis.

4.1 Participants

For the purpose of the study, initially 66 Iranian English-major B.A. students, comprising two intact "Listening and Speaking" classes, were selected as potential participants. However, some of the students were not at the intermediate level of proficiency as indicated by their Quick Placement Test (QPT) scores (N=5), failed to attend one or more treatment sessions (N=5), or failed to take the immediate or delayed post-tests (N=2). Therefore, from the original pool of 66 EFL learners, the final analyses were carried out on data obtained from 54 participants, comprising an individual output [IO] group (N=26) and a collaborative output [CO] group (N=28). They ranged in age from 19 to 31, with an average of 22.9.

4.2 Instruments

Two instruments were used in the present study: the Quick Placement Test (QPT), and a Written Discourse Completion Test (WDCT). Details of these measures are presented in this section.

Quick Placement Test (QPT): In order to homogenize the participants in terms of their general language proficiency, the paper-and-pencil version of the Quick Placement Test, developed by Oxford University Press and the Cambridge ESOL Examination Syndicate, was used. The test is comprised of 60 multiple-choice vocabulary, grammar, and cloze items. The results are reported along ALTE's scale: *Beginner* (0-10), *Breakthrough* (11-17), *Elementary* (18-29), *Lower Intermediate* (30-39), *Upper Intermediate* (40-47), *Advanced* (48-54), and *Very Advanced* (55-60). The first version of the test was used to identify and include only intermediate learners (i.e. those scoring between 30 and 47) in

the study. The complete test took about 30 minutes to complete. Moreover, the internal consistency of the test was acceptable, as indicated by a Cronbach's Alpha coefficient of .83.

Written Discourse Completion Test (WDCT): The pragmatic proficiency of the participants was measured through a 24-item Written Discourse Completion Test, comprising 8 situation prompts on each of the three speech acts of "apology," "request," and "refusal." The prompts reflected plausible situations in the life of university students and represented sufficiently varied combinations of the three social context variables of "power," "distance," and "imposition," following Brown and Levinson (1987). Responses were rated on a 6-point Likert scale developed by Taguchi (2006), which places a premium on three aspects of speech act performance: situational appropriateness, grammatical soundness, and discoursal felicity. The WDCT took about 50 minutes to complete. The test was also shown to have acceptable internal consistency, as indicated by a Cronbach's Alpha coefficient of .77 based on the participants' pretest scores.

4.3 Data Collection Procedure

The pre-treatment phase of the study involved two activities. The first one was to present parts of the two series "Lost" and "Friends," and the movie "Doubt," from which video input to IO and CO groups had been extracted, followed by class discussions of their themes and characters. The second one was to administer the QPT and the WDCT to the two experimental groups over two consecutive sessions. Based on the QPT scores, only intermediate learners were included in data analysis. This step was taken as a precaution against the possible match of the participants' grammatical and pragmatic competencies. However, the general research finding is that "a learner of high grammatical proficiency will not necessarily possess concomitant pragmatic competence" (Bardovi-Harlig, 2001, p. 14). As for the treatment phase, the IO and CO treatments were distributed over nine weekly sessions, divided into three 3-week sections allocated to "apologies," "requests," and "refusals," respectively. Finally, at the post-treatment phase, the participants took the WDCT twice: one week

after the last treatment session as the immediate posttest and four weeks after the last treatment session as the delayed posttest.

Instructional materials: The IO and CO groups were exposed to input in the form of 30 speech act-contained video excerpts during the treatment period, 10 for each of the three speech acts. The excerpts were extracted from different episodes of two popular series: "Lost" and "Friends," and from the movie "Doubt." The themes of the speech situations resembled those in the real life of the participants. The situations represented various degrees of weightiness and various semantic formulae. The 30 selected excerpts (23 from the series "Lost," 4 from the series "Friends," and 3 from the movie "Doubt") covered such role relationships as close friends, colleagues, teacher□student, teacher□school principal, distant acquaintances of the same or different age(s), mother□son, father□son, doctor□patient, etc. They also varied in length from 10 seconds (10□) to 2 minutes and 20 seconds (2□.20□).

Instructional treatments: The two output conditions were implemented in a way to offer ample pushed output opportunities during the three sessions dealing with each speech act. Identical output tasks made up the IO and CO conditions, except for the fact that IO tasks were completed individually, but CO tasks were completed in same-gender pairs (4 female and 10 male pairs), with the teacher providing no metapragmatic information to either group. As for the CO group, pairs were formed based on an informal survey of the participants' own preferred partners. Moreover, in order to make sure that the completion of all the tasks involved real peer collaboration (i.e. the collaborative discussion of pragmalinguistic and sociopragmatic features of the speech situation), the teacher (one of the researchers) orally modeled each of the tasks with a student in advance of student task performance.

Two main types of output tasks were designed: (1) output production tasks and (2) output manipulation tasks. The distinction between output production and manipulation tasks has been referred to by Ellis (2008), who views output-based instruction as a major methodological option in instructed SLA. The instructional procedure was consistent across the three speech acts for both

groups. The following is a speech act-independent account of the procedure implemented over a three-session period:

Session One

(A) Presentation of the first four video excerpts and their transcripts one by one;
(B) A "personalizing" (warm-up) phase, involving learners in recalling their speech act related experiences in situations similar to those in the video excerpts to which they were exposed;
(C) An output production task in which learners were presented with the transcripts of four more video excerpts from the same series and movie, with the speech act statements left out and a brief situation description appearing on top of each; subsequently, they were asked to fill in the blanks with the speech act in question. Upon task completion, they were presented with the excerpts and their original transcripts.

Session Two

(A) Presentation of the next three video excerpts and their transcripts one by one;
(B) An output production task in which learners completed five WDCTs on the speech act in question; and
(C) An output manipulation task in which learners were asked to exchange their output in phase B with another student's or pair's, think of the ways in which their peers' output could be improved, and put down their recommended modifications.

Session Three

(A) Presentation of the last three video excerpts and their transcripts one by one;
(B) An output production task in which learners were asked to write two 4-to-8-turn conversations, based on two situation prompts requiring them to make the speech act in question; and
(C) An output manipulation task in which learners were asked to exchange their output in Phase B with another student's or pair's, see whether their peers' output could be improved, and put down their recommended modifications.

4.4 Data Analysis

Prior to the treatment period, Cronbach's Alpha coefficients were calculated to test the internal consistencies of the QPT and the WDCT. IO and CO effects were compared through a mixed between-within groups ANOVA. Subsequently, two separate repeated-measures ANOVAs were conducted to provide separate accounts of IO and CO's short-term and long-term effects. In addition, three independent samples *t* tests were conducted to compare IO and CO groups in terms of their pretest, immediate posttest, and delayed posttest WDCT scores.

5. Results

In order to compare the short-term and long-term effects of IO and CO on the participants' WDCT scores, a mixed ANOVA was run. First, four preliminary assumptions were checked: (1) Ratios of skewness and kurtosis to their associated standard errors fell within the range of $\pm$1.96 for all the six score sets, indicating normality at the .05 level of significance (see Table 1 for descriptive statistics); (2) an insignificant Box's M statistic showed equality of intercorrelations [*Box's M*=4.653, *p*=0.628]; (3) a significant Mauchly's test statistic indicated violation of the assumption of equality of variances [$\chi^2(2)$=.796, *p*<.05], which necessitated the use of the conservative Greenhouse-Geisser's test; and finally (4) an insignificant *t* statistic obtained in an independent samples *t* test conducted on IO and CO groups' pretest WDCT scores indicated no significant mean difference [M_{IO}=3.064, M_{CO}=3.075, *t(52)*=.211, *p*>.05].

Table 1: Descriptive statistics for IO and CO groups' WDCT scores

Group	WDCT	Mean	SD	Skewness		Kurtosis	
				Statistic	SE	Statistic	SE
IO	Pretest	3.06	.21	.14	.45	-.46	.88
	Immediate	3.60	.20	.25	.45	-.91	.88
	Delayed	3.72	.17	.24	.45	-.22	.88
CO	Pretest	3.07	.19	-.55	.44	-1.224	.85
	Immediate	3.97	.17	-.00	.44	-.044	.85
	Delayed	4.17	.18	-.30	.44	-.206	.85

Note: IO= individual output, CO= collaborative output.

Subsequent to checking the preliminary assumptions, a mixed between-within subjects ANOVA was conducted. The results showed a significant main effect for Time [F=1026.338, p<.05, partial eta squared=.952], Group [F=36.193, p<.05, partial eta squared=.410], and Time and Group in combination [F=62.519, p<.05, partial eta squared=.546] (see Table 2).

Table 3 presents the pairwise mean comparisons for the effects of Time and Group, separately. As for the effect of Time, the results showed the participants' statistically significant gains not only form the pretest to the immediate posttest [*Mean difference*=.721, p<.05] and from the pretest to the delayed posttest [*Mean difference*=.883, p<.05], but also from the immediate posttest to the delayed posttest [*Mean difference*=.162, p<.05]. Regarding the effect of Group, the CO group's mean score was significantly higher than that of the IO group [*Mean difference*=.276, p<.05]. Moreover, the interactive effect of Time and Group showed that the IO and CO groups interacted differently with Time (or tests over time) regarding their WDCT performance.

Table 2: ANOVA results for the effects of Time, Group, and Time*Group on WDCT scores

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Eta^2
Time	Sphericity Assumed	23.83	2	11.91	1026.33*	.000	.952
	Greenhouse-Geisser	23.83	1.66	14.35	1026.33*	.000	.952
	Huynh-Feldt	23.83	1.74	13.68	1026.33*	.000	.952
	Lower-bound	23.83	1.00	23.83	1026.33*	.000	.952
Time * Group	Sphericity Assumed	1.45	2	.72	62.51*	.000	.546
	Greenhouse-Geisser	1.45	1.66	.87	62.51*	.000	.546
	Huynh-Feldt	1.45	1.74	.83	62.51*	.000	.546
	Lower-bound	1.45	1.00	1.45	62.51*	.000	.546
Error (Time)	Sphericity Assumed	1.20	104	.01			
	Greenhouse-Geisser	1.20	86.35	.01			
	Huynh-Feldt	1.20	90.57	.01			
	Lower-bound	1.20	52.00	.02			
Group	Intercept	700.71	1	700.71	24752.6*	.000	.998
	Groups	1.02	1	1.02	36.19*	.000	.410
	Error	1.47	52	.02			

Note: IO= individual output, CO= collaborative output.
*. The F-ratio is significant at the .05 level.

Table 3: Pairwise mean comparisons for the effects of Time and Group

Source		Mean Difference	SE	Sig.
Time	Pretest to Immediate posttest	-.721*	.021	.000
	Pretest to Delayed posttest	-.883*	.024	.000
	Immediate posttest to Delayed posttest	-.162*	.016	.000
Group	IO-CO	-.276*	.046	.000

Note: IO= individual output, CO= collaborative output.
*. The mean difference is significant at the .05 level.

Figure 1 illustrates the groups' patterns of performance, plotted on the basis of their pretest, immediate posttest, and delayed posttest WDCT mean scores.

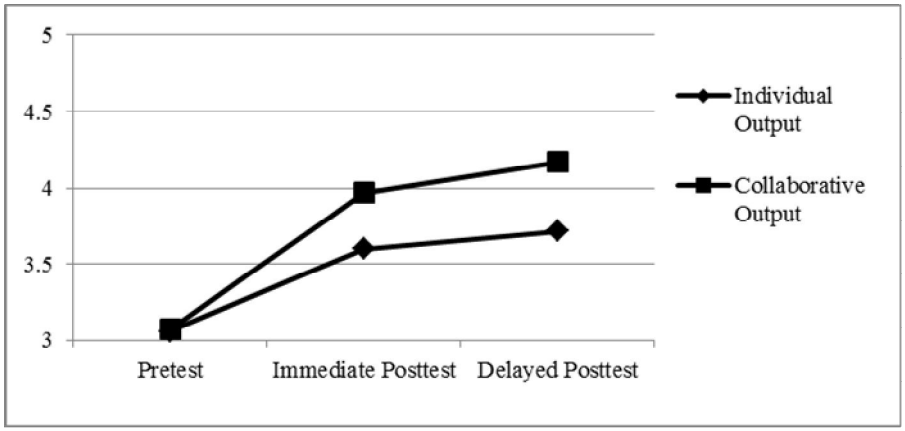

Figure 1: Patterns of WDCT performance of IO and CO groups

In order to test the statistical significance of the two groups' pragmatic improvements from the pretest to the immediate and delayed posttests and from the immediate posttest to the delayed posttest, two separate repeated measures ANOVAs were conducted, one for each group (see Table 1 for descriptive statistics). Regarding the IO group, Mauchly's test indicated that the assumption of sphericity had been met [$\chi^2(2)=3.840$, $p>.05$]. As Table 4 shows, there was a statistically significant effect for Time as the within-subject variable, with a large effect size [$F(2, 24)= 340.259$, $p<.05$, partial eta squared=.932] for the IO group.

Table 4: ANOVA results for the individual output group's WDCT scores

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Eta^2
Test	Sphericity Assumed	6.52	2	3.26	340.25*	.000	.93
	Lower-bound	6.52	1.00	6.52	340.25*	.000	.93
Error (Test)	Sphericity Assumed	.47	50	.010			
	Lower-bound	.47	25.00	.01			

Note: *. The F-ratio is significant at the .05 level.

Post hoc mean comparisons showed that the IO group made statistically significant gains (1) from the pretest to the immediate posttest [*Mean difference*=.543, $p<.05$], (2) from the immediate posttest to the delayed posttest [*Mean difference*=.122, $p<.05$], and consequently (3) from the pretest to the delayed posttest [*Mean difference*=.665, $p<.05$]. The gain score from the pretest to the immediate posttest was much larger than that from the immediate posttest to the delayed posttest.

With respect to the immediate and delayed effects of CO on WDCT scores, Mauchly's test produced a statistically significant statistic, which is an indication of the violation of the sphericity assumption [$\chi^2(2)=8.929$, $p<.05$]. Table 5 presents the results of the repeated measures ANOVA for the CO group. With sphericity not assumed, even the Greenhouse-Geisser correction yielded a statistically significant *F*-statistic, with a large effect size [$F=712.918$, $p<.05$, partial eta squared=.964]. The conclusion follows that Time had a significant main effect on the CO group's WDCT performance.

Table 5: ANOVA results for the collaborative output group's WDCT scores

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Eta^2
Test	Sphericity Assumed	19.24	2	9.62	712.91*	.000	.96
	Greenhouse-Geisser	19.24	1.55	12.41	712.91*	.000	.96
	Huynh-Feldt	19.24	1.62	11.83	712.91*	.000	.96

	Lower-bound	19.24	1.00	19.24	712.91*	.000	.96
Error (Test)	Sphericity Assumed	.72	54	.01			
	Greenhouse-Geisser	.72	41.83	.01			
	Huynh-Feldt	.72	43.90	.01			
	Lower-bound	.72	27.00	.02			

Note: *. The F-ratio is significant at the .05 level.

Based on the results of post hoc pairwise mean comparisons, the CO group made statistically significant gains (1) from the pretest to the immediate posttest [*Mean difference*=.899, $p<.05$] , (2) from the immediate posttest to the delayed posttest [*Mean difference*=.202, $p<.05$], and (3) from the pretest to the delayed posttest [*Mean difference*=1.101, $p<.05$]. Moreover, the participants' improvement in speech act production was more marked from the pretest to the immediate posttest.

Subsequent to testing the statistical significance of the short-term and long-term IO and CO effects, two separate independent samples *t* tests were run on IO and CO groups' (1) immediate posttest WDCT scores and (2) delayed posttest WDCT scores (see Table 6). The results showed that despite no significant pretreatment difference, the CO group significantly outperformed the IO group on both the immediate posttest [$t(52)$=-7.067, $p<.05$, partial eta squared=.48] and the delayed posttest [$t(52)$=-9.553, $p<.05$, partial eta squared=.63]. Moreover, CO's gain score from the immediate posttest to the delayed posttest (Mean difference_{CO}=.202) was greater than IO's (Mean difference_{IO}=.122).

Table 6: Independent samples *t* test results for individual output (IO) and collaborative output (CO) groups' immediate and delayed WDCT scores

WDCT	Levene's Test		T-test for Equality of Means			
	F	Sig.	t	df	Sig.	Mean Difference
Immediate	1.95	.16	-7.06*	52	.000	-.36
Delayed	.05	.82	-9.55*	52	.000	-.44

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

In summary, both groups made significant immediate and delayed gains in their speech act production ability, but the CO group was superior in both cases.

6. Discussion

The present study substantiated the short-term and long-term effects of individual and collaborative output-based instruction on EFL learners' speech act production ability, with the latter exerting a greater positive short-term and long-term influence. WDCT mean score differences reached statistical significance from the pretest to the immediate and delayed posttests for both IO and CO groups. However, partial eta squared values produced in the two ANOVAs run for the two groups were very large exceeding .9. Based on Coe's (2002) discussion of factors influencing effect size, the possibly overestimated effect size values are explicable considering (1) the small sample size and (2) the small standard deviations of pretest, immediate posttest, and delayed posttest WDCT scores, falling in excess of .218 for the IO group and .192 for the CO group.

That IO and CO conditions positively influenced the participants' speech act production ability can be taken as an indication of the potential of learner output, produced either individually or collaboratively, for ILP development. Both types of output might have engaged the participants in one or more of Swain's (1985, 1995, 2005, 2006) postulated functions of learner output: (1) noticing pragmatic knowledge gaps and relevant pragmatic features in the subsequent input, (2) reflection over pragmalinguistic and sociopragmatic aspects of speech act production, and (3) hypothesis generation and testing as to the appropriate production of the speech act at issue. This is probably the case, since both groups made significant improvements even from the immediate to the delayed posttest.

As for the effect of individual output, the participants probably received internal feedback (i.e. own feedback) or auto-input on their output productions and manipulations. Moreover, the participants were exposed to tailor-made input (i.e. video input containing the targeted speech acts) before and after the individual output task on

the first treatment session. The output task might have primed them to notice gaps in their speech act production ability, hence their enhanced noticing of and reflection over the relevant pragmalinguistic and sociopragmatic features of subsequently supplied input (Swain, 1985, 1995). Swain and Lapkin (1995) force the issue: "Output sets 'noticing' in train, triggering mental processes that lead to modified output" (p. 371). The observed positive impact echoes Izumi's (2002) finding as to the positive influence of individual output, following an input exposure phase, on the acquisition of English relativization; however, his main output task was an individual text reconstruction one, and the study was carried out in an ESL context. The observed effect might also be due to Jernigan's (2007) finding as to the positive impact of individual output on learners' perceptions of pragmatics and speech act appropriateness. Jernigan substantiated a significant main effect for individual output in terms of only expressive speech acts, which he attributed to the deeper processing learner output entails, but not directive speech acts, which he accounted for by speculating a possible ceiling effect in his participants' speech act production ability. The present study, however, did not involve investigating the differential effects of individual output on different speech acts; rather, it substantiated a significant overall positive impact, irrespective of the types of targeted speech acts. Moreover, while Izumi (2002) and Jernigan's (2007) studies involved a control (-output) group, the present study employed a comparison group design. Accordingly, any remarks as to impact of individual output per se should be made cautiously. Finally, neither study involved the investigation of the long-term effect of individual output through a delayed posttest.

On the other hand, the main factors in CO group's improved performance on both immediate and delayed posttests might have been (1) the external feedback they received from their peers in output production tasks, and (2) the collaborative/dialogic reflections and metalinguistic talk they obligatorily engaged in while carrying out the output manipulation tasks. In other words, the CO group likely experienced "languaging," rather than just "output." Instances of knowledge gap noticing, reflection, and

hypothesis testing by the CO group must have been discussed in greater depth, or otherwise resolved, on account of the availability of both internal and external feedback. This has been shown in the qualitative analyses of individually and collaboratively produced output in comparative studies of individual/collaborative output (e.g., Nassaji & Tian, 2010; Storch, 2005, 2007). The finding as to the positive impact of collaborative output on ILP development is at odds with Cohen's (2009) speculation that collaborative output in peer-peer dyads cannot induce ILP development. In his elaboration of the concept of "comprehensible pragmatics," Cohen (2009) states that peer-peer dialoging might not be as knowledge-conducive in pragmatics as in other aspects of communicative competence, owing to the implication of sociocultural values in pragmatic performance. If, he instantiates the issue, peers share the same linguistic and sociocultural background, which is the case in EFL contexts, non-targetlike pragmatic production might seem quite acceptable in terms of politeness, formality, and appropriateness, and this deteriorates the necessary conditions for output modification, and consequently for ILP development. The results of the present study provide counterevidence to this claim: Same language and sociocultural background does not seem to be the main factor in determining collaborative dialoging's potential for ILP development; rather, the depth of consciousness of relevant pragmatic features that collaborative output raises in learners, as evident in its lingering positive effect, can play a determining role. It needs to be admitted that the two groups' improved performance on the delayed posttest might have been in part due to the *test* effect owing to the triple administration of the same test, which, if true, should have influenced both groups.

The present study also showed the greater short-term and long-term benefits of collaborative output. It seems likely that collaborative output tasks are more fulfilling in terms of the generation of meta-talk, or "language-related episodes" in Swain and Lapkin's (1995) words, and output modification opportunities. Moreover, the probability of correct problem solution in collaborative dialoging seems to be higher, owing to the availability of peer feedback as well as the possibility of dyadic knowledge

pooling (Swain, 2000, 2006). The finding as to the superiority of collaborative over individual output is in agreement with Jabbarpoor and Tajeddin's (2013) finding, though their study addressed the acquisition of the English subjunctive mood. However, some studies have shown no significant difference between individual and collaborative output tasks (e.g., Nassaji & Tian, 2010; Storch, 2005). With respect to the present study's contradictory findings, it seems likely that the acquisitional relevance of output tasks, whether individual or collaborative, is partly a function of the targeted language features. Interlanguage pragmatics might benefit from collaborative dialoging more than grammar and vocabulary. Moreover, the modeling of paired interaction by the teacher at the pre-treatment phase of the present study might have contributed to the enrichment of collaborations, and subsequently to the superior performance of the collaborative output group. In other studies, "the interactions may not have been rich enough to lead to the appropriation and internalization of the word knowledge" (Nassaji & Tian, 2010, p. 412). Overall, given the few comparative studies of individual and collaborative output, there is need for more empirical research before making any definitive conclusions.

7. Conclusion

As an important aspect of SLA research, ILP development can be particularly challenging in EFL settings for many reasons (Liu, 2007; Tello Rueda, 2004). One major reason related to this study is the limited opportunity for language use. In this context, what learners themselves can contribute to the process of their interlanguage pragmatic development gains an increasing salience. Among learner-related factors "learner output" stands out, but its conspicuous under-representation in ILP research and practice is highly regrettable. The results of the present study revealed that the postulated benefits undergirding learner output produced either individually or collaboratively apply to the realm of ILP development. Pushed output, in the absence of implicit or explicit teacher feedback, showed significant short-term and long-term

promise for enhancing EFL learners' speech act production ability, though CO tasks turned out to better fulfill this potential.

These findings warrant consideration of output-based instruction alongside merely input-based instructional approaches, which have overwhelmingly dominated interventionist ILP research. Individual/collaborative output, when elevated to the core of classroom practices, with tailor-made input provided only for the sake of output task completion, would be an effective instructional pragmatics approach. Such an approach presumably has the potential to create in learners the urge for continuous ILP development, as shown in the significant improvement of both groups from the immediate posttest to the delayed posttest in the present study. The superior performance of the CO group probably indicates the significance of peer feedback, metalinguistic talk, and the cognitive, communicative and social activity CO engages learners in for ILP development (van Lier, 2000). Overall, it seems to be high time ILP practitioners abstracted away from the haunting dilemma of implicit or explicit pragmatic instruction, and propelled their practices into a consideration of learners' own potential, including their output production and modification.

The present study employed a comparison group design, addressed only the production of speech acts, and measured the participants' speech act production through a WDCT. Replication studies employing a control group design with a larger sample, targeting other aspects of L2 pragmatics (e.g. pragmatic routines, implicature, or speech act comprehension), and triangulating speech act data with other measures than the WDCT would function to confirm (or disconfirm) the study's findings.

References

Adams, R., Nuevo, A. M., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner-learner interactions? *The Modern Language Journal*, *95*(1), 42-63. doi:10.1111/j.15404781.2011.01242.x

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics? In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 13-32). Cambridge: Cambridge University Press.

Bardovi-Harlig, K., & Salsbury, T. (2004). The organization of turns in the disagreements of L2 learners: A longitudinal perspective. In D. Boxer & A. D. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 199-227). Bristol: Multilingual Matters.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage.* Cambridge: Cambridge University Press.

Coe, R. (2002). *It's the Effect Size, Stupid! What effect size is and why it is important.* Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England. Retrieved from www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, A. D. (2009). Comprehensible pragmatics: Where input and output come together. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 253-265). Poznan-Kalisz, Poland: Adam Mickiewics University Press.

Egi, T. (2010). Uptake, modified output, and learner perceptions of recasts: Learner responses as language awareness. *The Modern Language Journal*, *94*(1), 1-21. doi: 10.1111/j.1540-4781.2009.00980.x

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings, *Studies in Second Language Acquisition*, *21*(2), 319-333.

Fukuya, Y. J., & Zhang, Y. (2006). The Effects of recasting on the production of pragmalinguistic conventions of request by Chinese learners of English. *Issues in Applied Linguistics*, *15*(1), 59-91.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, *24*(4), 541-577. doi:10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, *34*(2), 239-287. doi: 10.2307/3587952

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*(3), 421-452.

Jabbarpoor, Sh., & Tajeddin, Z. (2013). Enhanced input, individual output, and collaborative output: Effects on the acquisition of the English subjunctive mood. *Revista Signos. Estudios de Lingüística*, *46*(82), 213-235. doi:10.4067/S0718-09342013000200003

Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 165-211). Amsterdam: John Benjamins.

Jernigan, J. E. (2007). *Instruction and developing second language pragmatic competence: An investigation into the efficacy of output* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3263885)

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

Kubota, M. (1995). Teachability of conversational implicature to Japanese EFL learners. *IRLT Bulletin*, *9*, 35-67.

Li, S. (2012). The effects of input-based practice on pragmatic development of requests in L2 Chinese. *Language Learning*, *62*(2), 403-438. doi: 10.1111/j.1467-9922.2011.00629.x

Liddicoat, A. J., & Crozet, C. (2001). Acquiring French interactional norms through instruction. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 125-144). Cambridge: Cambridge University Press.

Liu, J. (2007). Developing a pragmatics test for Chinese EFL learners. *Language Testing, 24*(1), 391-415. doi: 10.1177/0265532207077206

Mitchell, R., & Myles, F. (2004*). Second language learning theories* (2nd ed.). London: Hodder Arnold.

Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, *14*(4), 397-419. doi: 10.1177/1362168810375364

Rose, K. R., & Ng Kwai-fun, C. (2001). Inductive and deductive teaching of compliments and compliment responses. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 145-170). Cambridge: Cambridge University Press.

Shehadeh, A. (1999). Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning*, *49*(4), 627-675. doi: 10.1111/0023-8333.00104

Shehadeh, A. (2002). Comprehensible output, from occurrence to acquisition: An agenda for acquisitional research. *Language Learning*, *52*(3), 597-647. doi:10.1111/1467-9922.00196

Smith, B. (2005). The relationship between negotiated interaction, learner uptake and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly*, *39*(1), 33-58. doi: 10.2307/3588451

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, *14*(3), 153-173. doi: 10.1016/j.jslw.2005.05.002

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research*, *11*(2), 143-59. doi:1177/1362168807074600

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, Mass.: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and*

practice in applied linguistics: Studies in honour of H.G. Widdowson (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *The handbook of research in second language teaching and learning* (pp. 471-483). Mahwah, NJ: Lawrence Erlbaum.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced Language Learning: The contribution of Halliday and Vygotsky* (pp. 95-108). London: Continuum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371-391. doi:10.1093/applin/16.3.371

Taguchi, N. (2006). Analysis of appropriateness in a speech act of request in L2 Spanish. *Pragmatics*, *16*(4), 513-533.

Taguchi, N. (2011). Teaching pragmatics: Trends and issues. *Annual Review of Applied Linguistics*, *31*, 289-310. doi: 10.1017/S0267190511000018

Tajeddin, Z., Keshavarz, M. H., & Zand Moghadam, A. (2012). The effect of task-based language teaching on EFL learners' pragmatic production, metapragmatic awareness, and pragmatic self-assessment**.** *Iranian Journal of Applied Linguistics*, *15*(2), 139-166.

Takahashi, S. (2001). The role of input enhancement in developing pragmatic competence. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 171-199). Cambridge: Cambridge University Press.

Takahashi, S. (2010). Assessing learnability in second language pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 391-421). Berlin: Mouton de Gruyter.

Takimoto, M. (2007). The effects of input-based tasks on the development of learners' pragmatic proficiency. *Applied Linguistics*, *30*(1), 1-25. doi:10.1093/applin/amm049

Tello Rueda, L. Y. (2004). *Effects of pedagogical intervention on the development of pragmatic competence in adult learners of English as a foreign language (EFL)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3145136)

Van Lier, L. (2000). From input to affordance: Social interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245-259). Oxford: Oxford University Press.