Research in English Language Pedagogy (2025) 13(3): 130301

 DOI: 10.71673/relp.2025.1196041

Original Research

The Effect of Collaborative Tasks and Corrective Feedback Types on Iranian EFL Learners' Recall and Production of Complex Structures

Mozhgan Ebrahimian[1], *Bahman Gorjian*[2*], *Zohreh Seifoori*[3]
[1]Ph.D. student of Department of English Language, SR, C, Islamic Azad University, Tehran, Iran
[2]Associate Professor, Department of English Language, Arv, C, Islamic Azad University, Abadan, Iran
[3]Associate Professor, Department of English Language, SR, C, Islamic Azad University, Tehran, Iran

Submission date: 09-01-2025 Acceptance date: 03-03-2025

Abstract

Complex structures, which consist of dependent and independent clauses, make it difficult for Iranian high school students to recognize their grammatical complexities. This study investigates the effect of collaborative tasks (i.e., co-practice task writers, corrective feedback providers, and evaluators) on the types of corrective feedback (i.e., teacher feedback vs. peer feedback) and their impact on EFL learners' recall and production of complex structures. A quasi-experimental design was adopted, involving three equal intact classes comprising 96 lower-intermediate students, selected through convenience sampling. A production pretest and posttest of complex structures, as well as recall pretests and posttests of complex structures, were implemented following a pilot study to validate the tests. A two-way multivariate analysis of covariance (MANCOVA) was run. Findings revealed that collaborative tasks (co-practice task writers, corrective feedback providers, and evaluators) have a positive effect on high school students' recall and production of complex structures. Additionally, teacher feedback is a significant factor in students' recall and production of complex structures. No significant interaction was observed between explicit instruction of collaborative tasks and types of feedback on learners' recall and production of complex structures. The results address several suggestions for EFL teachers, learners, and pedagogical practitioners.

Keywords: Collaborative tasks, Complex structures, Corrective feedback, English grammar

* Corresponding Author's E-mail: bahman.gorjian@iau.ac.ir

1. Introduction

In today's globalized era, EFL learning has become a must in human life. Living in a global village is associated with some complexities in coping, at least some of which require EFL learning (Lightbown & Spada, 2020). English grammar is an inseparable part of EFL learning. As a main dimension of the English language, English grammar becomes difficult in the case of some structures. Teaching English grammar is a primary concern for language teachers in the Iranian context, as it can serve as the foundation for other language skills (Rezapoor & Mohammadzadeh, 2024). The question of how learners learn grammar can be one of the most critical and controversial issues in EFL research. Moreover, effective grammar instruction in foreign language contexts requires appropriate exposure to the language's grammar, which is learned as formally as possible.

English grammar is a multi-layered phenomenon that consists of many structures, including complex ones (Seifoori & Fattahi, 2014). The unique structure of complex structures creates major challenges for EFL learners in learning English grammar. It leads EFL learners either to skip these patterns or avoid using them in the production of spoken and written language. One such challenge is fossilization, wherein EFL learners' production patterns become fixed or stabilized at a certain developmental stage, which might impede further progress in language acquisition (Lyster & Ranta, 1997). Fossilization may occur due to limited exposure to comprehensible input, insufficient practice opportunities, or ineffective feedback mechanisms.

Additionally, learners may struggle with transfer difficulties when attempting to integrate complex structures from the target language into their spoken or written discourse. First language (L1) interference can result in errors and inaccuracies in the use of complex syntactic structures. Learners must navigate features of their L1, which may pose challenges in achieving native-like proficiency in complex sentence production (Wistner et al., 2013).

Contextual factors such as instructional methodology, curriculum design, and language learning environments influence learners' opportunities for practicing and mastering complex structures. Communicative language teaching approaches that emphasize meaningful interaction, task-based activities, and opportunities for language output encourage learners to experiment with and produce complex linguistic forms in authentic communicative contexts (Wang, 2024). Conversely, traditional grammar-focused

instruction that prioritizes rote memorization and rule-based learning may hinder learners' ability to internalize and use complex structures in meaningful communication. In other words, learning complex structures in the English language is a complex process influenced by learners' proficiency level, cognitive abilities, language input, and instructional practices.

2. Literature Review

Despite the significance of language proficiency development, learners encounter various challenges in acquiring and using complex linguistic forms effectively. Addressing these challenges requires a holistic approach that integrates principles of input enhancement, task-based instruction, and meaningful interaction to provide learners with opportunities for meaningful practice and mastery of complex structures in the target language. Grammatical complexity may be due to the combinations of two or more embedded clauses that are dependent or independent and should be linguistically and semantically interrelated. Another major difficulty may be due to the grammatical patterns that are different between the two languages. Grammatical complexity has been characterized as the combination of production units or minimal terminal units (t-units) (Seifoori & Fattahi, 2014).

In the context of Iran, mastering complex structures is one major difficulty for many students (Farhangi & Pourmohammadi, 2018). Differences between the structures of complex structures in Persian and English potentially contribute to the difficulty of teaching them for Iranian English learners. This is why several researchers (e.g., Farhangi & Pourmohammadi, 2018; Kourang Beheshti & Sadeghi, 2019; Pakseresht Mogharab et al., 2014) have proposed different methods to enhance Iranian students' learning of complex structures. Among English complex structures, relative clauses (RCs) and conditional sentences are particularly challenging for EFL learners (Abdolmanafi & Seifi, 2014; Marefat & Abdollahnejad, 2014; Rezai, 2011).

RCs and conditional sentences (i.e., if-clauses or if-Cs) are structures that introduce complexity, and students need to receive instruction about these structures to learn. However, RCs and if-Cs must be learned as universal linguistic phenomena that are of high importance in pedagogical studies (Abdolmanafi & Seifi, 2014).As put by Marefat and Abdollahnejad (2014), the syntactic nature of RCs makes them complex for EFL learners.

Relative clauses in English make a sentence structure where an embedded clause is used in another sentence with a co-referential noun/noun phrase or noun clauses. The relative clause is dependent and acts as an adjective within a main clause.

A conditional sentence with 'if' as a linguistic marker joins two clauses and shows a possible condition, while the second part of the sentence describes the consequence of the action. These sentences are called If-Clauses as well.

The literature (e.g., Kouhsarian et al., 2023) has shown that rare studies investigate the effects of collaborative tasks, corrective feedback, and complex structures in a single Iranian research context. There are few studies on the types of collaborative tasks and feedback types that affect students' achievement of complex structures. For instance, Javan Amani et al (2024) investigated the impact of collaborative tasks on students' complex structures under a general term. However, the types of collaborative tasks, including co-practice task writers, corrective feedback providers, and evaluators, have not come into focus. Thus, to fill this gap, the current research aims to uncover the potential and practical usability of collaborative tasks and corrective feedback types (i.e., teacher feedback vs. peer feedback) on EFL learners' complex structures. The research questions of the study are formulated as follows:

1. How do collaborative tasks (co-practice task writer, corrective feedback providers, and evaluators) significantly affect learners' recall/production of complex structures?
2. Which types of feedback (teacher feedback vs. peer feedback) significantly affect learners' recall/production of complex structures?
3. Is there any significant interaction between collaborative tasks and types of feedback on learners' recall/production of complex structures?

3. Methodology

3.1. Design and Context of the Study

In line with the research objective, a quasi-experimental design includes pre- and post-tests, as well as an intervention (Creswell, 2020). The participants in the study were 105 female high school students in the 10th grade at a high school in Tehran, Iran. They were selected through convenience sampling and took a placement test, the Oxford Placement Test (2001). Then, they were assigned to three experimental groups: the co-practice task

writer group (CPG), the corrective feedback providers group (CFG), and the evaluators group (EG).

3.2. Participants

The participants included three intact classes of lower-intermediate female high school students, who were selected through convenience sampling (Creswell, 2020). They were in the 10th grade. Each class contained 32 students. To homogenize the participants, those whose scores were between 30 and 39 on the Oxford Placement Test (OPT, 2001) were selected. Finally, 32 learners in each class (totalling 96 learners) remained as the main participants in the quantitative phase of the study. The three classes were assigned to three experimental groups (i.e., co-practice task writer group (CPG), corrective feedback providers group (CFG), and evaluators group (EG). They were between 15 and 17 years old (M = 16.23, SD = 0.61). They spoke Persian as their mother tongue. To ensure adherence to research ethics, participants were informed of the study's objectives. Moreover, their data anonymity and confidentiality were promised.

3.3. Instruments

The study employed four research instruments to collect data. The first instrument was a placement test for OPT, used to homogenize the research sample. It comprised 60 items covering reading (20 items), grammar (20 items), and vocabulary (20 items). The individuals who scored between 30 and 39, indicating lower-intermediate proficiency, were selected as participants in the study. The OPT is a standard test whose reliability and validity have been established as reported by Wistner et al. (2013).

The second instrument was a production pretest of complex structures, which consisted of 90 items, including correct uses of the verbs that were given in parentheses. Forty-five items were on conditional sentences (i.e., type I and II), and 45 items were on relative clauses. These items were included in tests randomly selected from the test bank of education offices. The reliability and validity of the pretest were already verified by the test designers in the test bank; however, the content validity of the test was assessed through expert judgment, as three ELT experts were asked to comment on the content validity of the test. Moreover, the reliability of the production pretest was calculated using the KR-21 formula (r = .91) after conducting a pilot study on ten lower-intermediate

learners, excluding the participants. The same procedure was conducted to assess the validity and reliability of both the production posttest and the recall pretest and posttest.

A production posttest of complex structures was developed by changing the format of the production pretest items, including the order of items and choices, to prevent learners from recalling the content of the pretest. The validity of the posttests was verified by having three ELT experts review the items and provide comments on each one. Then, the researcher modified or discarded several items based on the experts' suggestions. Finally, the reliability of the production posttest was calculated through the KR-21 formula of (r=.91) after conducting the pilot study on ten lower-intermediate learners other than the participants.

A recall pretest of complex structures was implemented to evaluate participants' recall of these structures. It consisted of 90 multiple-choice items. Forty-five items were multiple-choice items of type I and II conditional sentences. The same process was used for taking relative pronoun items. The items were randomly selected from the test bank of education offices. The authorities already corroborated the reliability and validity of the tests in the test bank. However, a pilot study was conducted to validate the test, and three ELT experts provided feedback on the test's appropriateness. The reliability of the test was calculated using the KR-21 formula (r=.87).

A recall posttest of complex structures was developed by modifying the format of the pretest items, including the order of items or choices, to prevent learners from recalling the pretest. The reliability and validity of the tests were estimated in a pilot study. The KR-21 reliability index of the recall posttest was calculated to be 0.96. Three ELT experts reviewed the content of the posttest, and after several modifications, they approved the content validity of the test.

3.4. Data Collection Procedure

Data collection processes, which included research sample selection, homogenization, and the establishment of three experimental groups (i.e., CPG, CFG, and EG), were followed. Then, the production and recall pretests were administered on two separate days to accommodate the participants' convenience. Thereafter, the treatment sessions began, during which all three groups were exposed to 12 sessions, each held for 90 minutes twice

a week. The first educational session was devoted to explaining the content and procedure of the treatment period to the participants.

In the CPG, six sessions were devoted to conditional sentences, and the subsequent six sessions to relative clauses. In the first educational session on conditional sentences, the teacher first taught the general structure of conditional sentences (types I and II) deductively. Then, the learners were divided into four groups, each containing eight members who would serve as co-practice task writers. The teacher wrote five multiple-choice items of conditional sentences on the whiteboard in each session for the groups to answer through group work. The items were randomly selected from the archives of the education offices. The members were asked by the teacher to collaboratively share their grammatical knowledge of conditional sentences, working in teams to help each other in the classroom (Wang, 2024). Peers followed a problem-solving approach to provide feedback to each other in either oral or written form. All the learners in the teams were asked to participate in the task.

In the CFG, six sessions were assigned to conditional sentences, and the subsequent six sessions to relative clauses. Within the first session of teaching conditional sentences, the teacher first presented the general structure of conditional sentences (types I and II) deductively. Then, the learners were divided into four groups, each consisting of eight members who were to serve as corrective feedback providers. That is, the group members were asked to help each other arrive at the correct answers. In each group, one learner was selected by the teacher as the head, who was responsible for helping others and playing the tutor role within the group. The teacher monitored group responses as a final resort. The teacher wrote five multiple-choice items of conditional sentences on the whiteboard in each session for the groups to answer through teamwork. Four groups were just provided with peer feedback in the absence of teacher feedback. In these two groups, the teacher just monitored the final answers of the groups. The same procedure was followed in the sessions devoted to relative clauses. The primary difference between CPG and CFG lies in the amount of feedback received from peers. In the former, learners shared their ideas and collaboratively worked together in the classroom to solve the problem through co-practice. However, in the latter, learners worked individually as a team to solve the problem and provided corrective feedback to one another. Finally, the teacher monitored the group's responses and provided corrective feedback as a last resort.

Finally, in EG, the first session was devoted to teaching conditional sentences; the teacher first taught the general structure of conditional sentences (type I and type II) deductively. The teacher wrote five multiple-choice items of conditional sentences on the whiteboard in each session for the students who were called evaluators to answer. The evaluators were asked to give oral or written feedback on each other's answers in paired groups under the supervision of the teacher. The pairs of learners worked together to solve the problem and evaluated each other's responses to the exercises. However, in CPG and CFG, learners worked either as a unified team or as individuals within a team that worked separately. The same procedure was followed in the sessions devoted to relative clauses. Finally, after the intervention sessions, the production and recall posttests were administered on two separate days to ensure the comfort of the participants.

3.5.Data Analysis Procedure

To analyze the data, descriptive statistics and two-way multivariate analysis of covariance (MANCOVA) were run. This statistical technique was employed to analyze the data and compare the mean scores of pretests and posttests of recall and production across groups, while controlling for the effect of pretests. Through the use of inferential statistics of Two-Way MANCOVA, the researcher successfully controlled for the influence of pretest scores, thereby ensuring that any differences noted in posttest scores could be linked to the experimental conditions instead of pre-existing disparities in the participants' abilities. This analysis facilitates a more detailed understanding of the interactions among various factors, as it takes into account multiple independent variables at once. Furthermore, it offered a solid framework for assessing the effects of the intervention on participants' recall and production skills, addressing the impact of pretest performance on the results. Thus, it can improve the validity of the findings and give a more precise interpretation of the outcomes.

4. Results

The research questions addressed the effect of collaborative tasks and types of feedback on learners' recall and production of complex structures. First, normality and Cronbach's alpha reliability tests were run, and the results are shown in Tables 1 and 2.

Table 1.
Normality Test

			Kolmogorov-Smirnov			Shapiro-Wilk		
Group	Feedback		Statistic	df	Sig.	Statistic	df	Sig.
CPG	Teacher	Pre-Recall	.155	16	.200	.932	16	.264
		Post Recall	.128	16	.200	.937	16	.310
		Pre-Production	.108	16	.200	.975	16	.906
		Post Production	.145	16	.200	.964	16	.737
	Peer	Pre-Recall	.211	16	.054	.948	16	.456
		Post Recall	.134	16	.200	.955	16	.581
		Pre-Production	.120	16	.200	.967	16	.788
		Post Production	.150	16	.200	.924	16	.195
CFG	Teacher	Pre-Recall	.203	16	.077	.898	16	.076
		Post Recall	.129	16	.200	.958	16	.623
		Pre-Production	.129	16	.200	.968	16	.813
		Post Production	.180	16	.174	.907	16	.103
	Peer	Pre-Recall	.130	16	.200	.942	16	.376
		Post Recall	.114	16	.200	.983	16	.982
		Pre-Production	.180	16	.175	.922	16	.181
		Post Production	.114	16	.200	.949	16	.468
EG	Teacher	Pre-Recall	.183	16	.157	.929	16	.238
		Post Recall	.133	16	.200	.929	16	.236
		Pre-Production	.174	16	.200	.962	16	.706
		Post Production	.176	16	.200	.933	16	.268
	Peer	Pre-Recall	.128	16	.200	.964	16	.742
		Post Recall	.190	16	.124	.918	16	.159
		Pre-Production	.164	16	.200	.968	16	.807
		Post Production	.155	16	.200	.947	16	.449

Table 1 illustrates that all mean scores are associated with indices that are higher than .05. Thus, the normality assumption was retained, and statistics of the KR-21 concerning recall and production posttests are represented in Table 2.

Table 2.
Indices of Test Reliability

	N	Mean	Std. Deviation	Variance	KR-21
Pre-Recall	96	43.69	8.734	76.280	.87
Post Recall	96	51.20	15.311	234.413	.96
Pre-Production	96	33.00	9.006	81.116	.79
Post Production	96	54.51	14.453	208.884	.91

Table 2 indicates that the reliability of pre/posttests of recall values are .87, and .96. For the writing production pre/posttests are .79 and .91 respectively It should be noted that Two-Way MANCOVA, besides the assumptions of normality and reliability which were reported in Table 1, and Table 2, has four more assumptions (i.e., homogeneity of variances of groups, linearity of relationships between pretests and posttests, homogeneity of regression slopes; and finally, homogeneity of covariance matrices). These assumptions are discussed below. The homogeneity of variances is displayed in Table 3.

Table 3.

Homogeneity Test of Variances for Posttests of Recall and Production

	F	df1	df2	Sig.
Post Recall	4.711	5	90	.001
Post Production	1.792	5	90	.123

Table 3 shows the homogeneity of variances in the writing production test (F (5, 90) = 1.79, p > .05); however, it was violated on the posttest of recall (F(5, 90) = 4.71, p < .05). If the groups enjoy equal sample sizes, as is the case in this study, the significant results of the Levene's test can be ignored. Table 4 presents the results of the assumption of linearity in the relationships between pretests and posttests of recall and production.

Table 4.

Tests of Linearity of Relationships between Pretests and Posttests of Recall and Production

			Sum of Squares	df	Mean Square	F	Sig.
Pre-Production × Post Production	Between Groups	(Combined)	4244.033	41	103.513	1.615	.049
		Linearity	2390.183	1	2390.183	37.282	.000
		Deviation from Linearity	1853.850	40	46.346	.723	.857
	Within Groups		3461.967	54	64.110		
	Total		7706.000	95			
	Eta Squared		.551				
Pre-Recall× Post Recall	Between Groups	(Combined)	4320.992	42	102.881	1.864	.016
		Linearity	1751.261	1	1751.261	31.725	.000
		Deviation from Linearity	2569.731	41	62.676	1.135	.329
	Within Groups		2925.633	53	55.201		
	Total		7246.625	95			
	Eta Squared		.596				

The linearity tests in Table 4 indicated that the assumption of linearity was retained on the production test (F (1, 95) = 37.28, $p < .05$, $\eta 2$ = .551 indicating a large effect size), and recall (F (1, 95) = 31.72, $p < .05$, $\eta 2$ = .596, representing a large effect size) tests.

Multivariate Analysis of Covariance requires that the linear relationships between the pre-tests and post-tests of recall and production are approximately consistent across the three groups, referring to the homogeneity of regression slopes.

Table 5.

Homogeneity Test of Regression Slopes for Recall and Writing Production Posttests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Group× Feedback × Pre-Recall × Pre-Production	Pillai's Trace	.118	.796	12	152	.654	.059
	Wilks' Lambda	.884	.793[b]	12	150	.658	.060
	Hotelling's Trace	.128	.789	12	148	.662	.060
	Roy's Largest Root	.098	1.236[c]	6	76	.297	.089

As indicated in Table 5, there is no significant interaction between the covariates (pretests) and the independent variable (types of collaborative tasks and feedback) (F (12, 152) = .796, $p > .05$, Partial $\eta 2$ = .059). This indicates a weak effect size. Thus, the statistical assumption that the relationships between pretests and posttests of recall and production were linear across groups was supported. Therefore, linear relationships were confirmed between the pretests and posttests of recall and production for the three groups.

Table 6 represents the Box's test of homogeneity of covariance matrices. Two-way MANCOVA requires correlations between any two pairs of dependent variables (i.e., posttests of recall and production) that are roughly equal across the groups.

Table 6.

Box's M Test of Equality of Covariance Matrices Posttests of Recall and Writing Production by Groups' Pretests

Box's M	39.644
F	2.494
df1	15
df2	44304.667
Sig.	.001

The Box's test (Box's M = 39.64, p = .001) depicts a non-significant result in Table 4.6. The assumption of homogeneity of covariance matrices was maintained. It is important to note that the results of Box's test should be significant at the .001 level. If the results of Table 6 are considered significant, the results of the Pillai's Trace test (Table 7) can be reported to compensate for the violation of this assumption. The test [Box's Test] evaluates the covariance matrices of dependent variables across different groups. A non-significant outcome indicates that the covariance matrices are approximately equal.

Table 7.

Multivariate Analysis of Covariance Posttests of Recall and Production by Groups with Pretests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.276	16.580	2.000	87.000	.000	.276
	Wilks' Lambda	.724	16.580	2.000	87.000	.000	.276
	Hotelling's Trace	.381	16.580	2.000	87.000	.000	.276
	Roy's Largest Root	.381	16.580	2.000	87.000	.000	.276
Pre-Recall	Pillai's Trace	.182	9.670	2.000	87.000	.000	.182
	Wilks' Lambda	.818	9.670	2.000	87.000	.000	.182
	Hotelling's Trace	.222	9.670	2.000	87.000	.000	.182
	Roy's Largest Root	.222	9.670	2.000	87.000	.000	.182
Pre-Production	Pillai's Trace	.435	33.549	2.000	87.000	.000	.435
	Wilks' Lambda	.565	33.549	2.000	87.000	.000	.435
	Hotelling's Trace	.771	33.549	2.000	87.000	.000	.435
	Roy's Largest Root	.771	33.549	2.000	87.000	.000	.435
Group	Pillai's Trace	.801	29.390	4.000	176.000	.000	.400

	Wilks' Lambda	.200	53.688	4.000	174.000	.000	.552
	Hotelling's Trace	3.985	85.684	4.000	172.000	.000	.666
	Roy's Largest Root	3.984	175.284	2.000	88.000	.000	.799
Feedback	Pillai's Trace	.301	18.775	2.000	87.000	.000	.301
	Wilks' Lambda	.699	18.775	2.000	87.000	.000	.301
	Hotelling's Trace	.432	18.775	2.000	87.000	.000	.301
	Roy's Largest Root	.432	18.775	2.000	87.000	.000	.301
Group×Feedback	Pillai's Trace	.050	1.133	4.000	176.000	.343	.025
	Wilks' Lambda	.950	1.133	4.000	174.000	.343	.025
	Hotelling's Trace	.053	1.132	4.000	172.000	.343	.026
	Roy's Largest Root	.051	2.244	2.000	88.000	.112	.049

The findings presented in Table 7 indicate notable differences among the overall means of three experimental groups regarding posttest performance in recall and writing production, after accounting for the influence of the pretest (F (4, 176) = 29.39, $p < .05$, Partial $\eta 2$ = .400, which signifies a large effect size). Additionally, a significant difference was observed between the overall means of the teacher feedback group and the peer feedback group in their posttest results for recall and production, also after controlling for the pretest effect (F (2, 87) = 18.77, $p < .05$, Partial $\eta 2$ = .301, indicating a large effect size). However, no significant interaction was found between the types of collaborative tasks and feedback regarding overall recall and production (F(4, 176) = 1.13, $p > .05$, Partial η^2 = .025, which reflects a weak effect size).

In cases where this assumption is violated, the Pillai–Bartlett Trace (Pillai's Trace) statistics in MANOVA are utilized. Therefore, the use of Pillai–Bartlett Trace (Pillai's Trace) statistics in MANOVA is advised due to their robustness. Consequently, Pillai–Bartlett Trace (Pillai's Trace) statistics in MANOVA are recommended for their robustness. After discussing the assumptions related to two-way MANCOVA, Table 7 reports the main results.

Table 8 illustrates the posttest means of recall for the three experimental groups, adjusted for the effects of the pretest.

Table 8.

Statistics for Posttests of Recall and Production by Groups with Pretests

Dependent Variable	Group	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Post Recall	CPG	50.985[a]	1.318	48.366	53.604
	CFG	64.010[a]	1.298	61.430	66.590
	EG	38.599[a]	1.307	36.001	41.197
Post Production	CPG	54.809[a]	1.174	52.476	57.142
	CFG	65.683[a]	1.157	63.385	67.981
	EG	43.039[a]	1.164	40.725	45.353

a. Covariates values: Pre-Recall = 43.69, Pre-Production = 33.00.

Table 8 shows that the CFG group achieves the highest means on the post-test of recall (*M* = 64.01, *SE* = 1.29). However, the CPG (*M* = 50.98, *SE* = 1.31) and EG (*M* = 38.59, *SE* = 1.30) groups are listed in the second and third rows. Table 9 shows Between-Subject Effects.

Table 9.

Posttests of Recall and Production Tests of (Between-Subjects Effects for Groups by Feedback with Pretests)

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-Recall	Post Recall	551.275	1	551.275	10.336	.002	.105
	Post Production	463.971	1	463.971	10.963	.001	.111
Pre-Production	Post Recall	1324.029	1	1324.029	24.825	.000	.220
	Post Production	2060.461	1	2060.461	48.684	.000	.356
Group	Post Recall	10182.029	2	5091.015	95.453	.000	.684
	Post Production	8080.610	2	4040.305	95.463	.000	.685
Feedback	Post Recall	1264.702	1	1264.702	23.712	.000	.212
	Post Production	744.788	1	744.788	17.598	.000	.167
Group× Feedback	Post Recall	205.623	2	102.812	1.928	.152	.042
	Post Production	22.527	2	11.263	.266	.767	.006
Error	Post Recall	4693.527	88	53.336			
	Post Production	3724.458	88	42.323			
Total	Post Recall	273907.00	96				
	Post Production	305097.00	96				

Table 9 shows significant differences between the CPG (*M* = 50.98), CFG (*M* = 64.01), and EG (*M* = 38.59) groups' posttest for recall, after accounting for the influence of the pretest, indicating a significant effect (*F* (2, 88) = 95.45, *p* < .05, Partial *η2* = .684). This suggests a large effect size. The subsequent post-hoc analysis is depicted in Table 10, showing recall in relation to the groups' pretest.

Table 10.
Pairwise Post-hoc Comparisons Tests for Posttest of Recall by Groups with Pretest

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
CFG	CPG	13.025*	1.863	.000	8.479	17.571
	EG	25.411*	1.840	.000	20.921	29.901
CPG	EG	12.386*	1.881	.000	7.795	16.977

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

Table 10 presents the results of the post-hoc comparison tests, indicating that the CFG group (*M* = 64.01) significantly outperforms the CPG group (*M* = 50.98) in the post-test of recall, with a mean difference of 13.02 (*p* < .05), after adjusting for the pretest effect.

The CFG group (*M* = 64.01) also demonstrated a significant advantage over the EG group (*M* = 38.59) in the post-test of recall, with a mean difference of 25.41 (*p* < .05), again after controlling for the pretest. Moreover, the CPG group (*M* = 50.98) significantly outperformed the EG group (*M* = 38.59) in the post-test of recall, with a mean difference of 12.38 (*p* < .05), after adjusting for the pretest effect. The post-hoc comparison tests for the post-test, taking the pre-test into account, are illustrated in Table 10.

Table 11.
Pairwise Post-hoc Comparisons Tests for Posttest of Production by Groups with Pretest

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
CFG	CPG	10.874*	1.659	.000	6.824	14.923
	EG	22.644*	1.639	.000	18.644	26.644
CPG	EG	11.770*	1.676	.000	7.680	15.860

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

Table 11 displays that the CFG group (M = 65.68) significantly outperformed the CPG group (M = 54.80) on the post-test of generation (MD = 10.87, p < .05) after controlling for the pretest effect. Additionally, it was found that the CFG group (M = 65.68) significantly outperformed the EG group (M = 43.03) on the posttest of generation (MD = 22.64, p < .05) after adjusting for the influence of the pretest. Lastly, the CPG group (M = 54.80) significantly outperformed the EG group (M = 43.03) on the posttest of generation (MD = 11.77, p < .05) after controlling for the pretest effect.

Table 12.
Statistics for Posttests of Recall and Production by Types of Feedback with Pretests

Dependent Variable	Feedback	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Post Recall	Teacher	54.87[a]	1.061	52.764	56.979
	Peer	47.52[a]	1.061	45.417	49.632
Post Production	Teacher	57.33[a]	.945	55.452	59.207
	Peer	51.69[a]	.945	49.814	53.569

a. Covariates appearing in the model are evaluated at the following values: Pre-Recall = 43.69, Pre-Production = 33.00.

Table 12 presents that the teacher feedback group (M = 54.87) had a significantly higher mean than the peer feedback group (M = 47.52) on the post-test of review after controlling for the pretest effect. It also displays the means of the teacher and peer feedback groups on the posttest of production after controlling for the influence of the pretest. Table 13 presents the descriptive statistics regarding the interaction between various types of collaborative tasks and feedback.

Table 13.
Descriptive Statistics for Interaction between Groups and Types of Feedback for Posttests of Recall and Production with Pretests

Dependent Variable	Group	Feedback	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Post Recall	CPG	Teacher	53.96[a]	1.849	50.285	57.634
		Peer	48.01[a]	1.851	44.331	51.689
	CFG	Teacher	69.83[a]	1.875	66.110	73.562
		Peer	58.18[a]	1.889	54.430	61.938
	EG	Teacher	40.81[a]	1.866	37.111	44.527
		Peer	36.37[a]	1.832	32.738	40.020

Post Production	CPG	Teacher	58.18[a]	1.647	54.913	61.460
		Peer	51.43[a]	1.649	48.155	54.709
	CFG	Teacher	67.83[a]	1.670	64.514	71.152
		Peer	63.53[a]	1.683	60.189	66.877
	EG	Teacher	45.96[a]	1.662	42.666	49.272
		Peer	40.10[a]	1.632	36.865	43.352

a. Covariate values: Pre-Recall = 43.69, Pre-production = 33.00.

Table 13 reveals no significant interaction between the explicit instruction of collaborative tasks and the feedback on the posttest of production. The CFG group achieves the highest mean scores for post-test writing production across both teacher and peer feedback groups, followed by the CPG and EG groups.

5. Discussion

The findings from the initial research question indicate that tasks involving the provision of corrective feedback were more effective than those involving co-practice writing. Additionally, co-practice writing tasks outperformed the evaluators' tasks in terms of enhancing learners' recall and production of complex structures. To what extent does explicit instruction of collaborative tasks (co-practice task writer, corrective feedback providers, and evaluators) significantly affect learners' recall and production of complex structures? Explicit instruction through collaborative tasks (co-practice task writer, corrective feedback providers, and evaluators) significantly affects students' recall and production of complex structures. Additionally, co-practice writing tasks are more effective than the other tasks, enhancing learners' recall and production of complex structures.This means that the most effective collaborative task in terms of learners' recall and production of complex structures was the provision of corrective feedback.

The results derived from the second research question, which examined the extent to which various forms of feedback (teacher feedback compared to peer feedback) significantly affect learners' ability to recall and produce complex structures, indicated that teacher feedback types do have a significant influence. Additionally, the findings illustrated that teacher feedback was more effective than peer feedback in improving learners' recall and production of complex structures.

The findings of this study highlight the significant impact of teacher and peer feedback on students' ability to recall and apply complex structures, which aligns with the

research conducted by Izadpanah et al. (2023). Their investigation revealed a notable influence of teacher feedback on the acquisition of grammatical rules among EFL learners. Moreover, these findings align with the research of Ghoorchaei et al. (2022), which investigated the effects of both teacher and peer corrective feedback on the retention of grammatical structures in the short and long term among Iranian EFL students. Moreover, the results support Sippel's (2021) study, which assessed the effectiveness of teacher and peer feedback, indicating a marked enhancement in the grammatical accuracy of EFL learners as a result of such feedback.

To account for the study's results, teacher and peer feedback promote ongoing reflection and metacognition, empowering students to monitor their learning progress and identify areas for growth (Braund & DeLuca, 2018). It contributes to learning, enhancing performance, and improving academic achievement. Teacher and peer feedback provide learners with clear direction and meaningful insights into their progress and areas for improvement. It leads to enhanced metacognitive awareness, self-regulation, and reflective thinking, allowing individuals to adapt their strategies based on feedback information. It enhances motivation, engagement, and self-efficacy, leading to improved performance and academic outcomes (Hattie & Timperley, 2007).

Teacher and peer feedback promote the recall and production of complex structures by fostering a growth mindset and providing opportunities to notice and internalize target language forms (Lyster & Ranta, 1997). Corrective feedback, regardless of its source, provides a range of perspectives and insights that can enhance students' understanding of their strengths and areas for improvement. This encourages students to consider alternative viewpoints, question assumptions, and reflect critically in the production phases (Van den Berg et al., 2008).

Teacher corrective feedback is highly effective in enhancing learners' recall and production of complex structures, and it may encourage learners to reflect on their language usage and accuracy (Sheen, 2004). It promotes learners' willingness to incorporate linguistic modifications into their speech production (Li, 2010). It also helps learners identify their strengths and weaknesses, think critically, communicate effectively, and engage in reflective practice (Carless & Boud, 2018).

Teacher feedback provides learners with expert guidance, direction in their language development (Ellis, 2009), explanations, and strategies for improvement, aligning with

Lightbown and Spada‘s (2020) notion that teacher feedback has a significant impact on enhancing learners' language achievement. This is confirmed by Ellis (2009), who states that helping students internalize correct language forms and grammatical structures is essential. Furthermore, effective teacher feedback promotes learners' awareness of language rules and conventions, contributing to learning (Kang et al., 2007). The timing and nature of teacher feedback are critical factors in its impact on learners' learning, aligning with Russell and Spada’s (2006) thoughts that teacher feedback enables students to make connections between their language production and the correct forms. Teacher feedback facilitates learning activities and enhances learners’ automaticity, aligning with Hattie and Timperley's (2007) assumptions that teacher corrective feedback should be specific and focus on linguistic accuracy to enhance learners' learning by addressing their immediate language needs. This result is also stated by several scholars (e.g., Ellis, 2009).

The findings about the research question one: ‘Is there any significant interaction between explicit instruction of collaborative tasks, and types of feedback on learners' recall and production of complex structures?’ proved that there was no significant interaction between explicit instruction of collaborative tasks and types of feedback on learners' recall and production of complex structures. Results indicate no significant interaction between explicit instruction of collaborative tasks and types of feedback on learners' recall and production of complex structures. This finding aligns with Pakseresht Mogharab et al. (2014), who reported no significant interaction between the explicit instruction of collaborative tasks and types of feedback on learners' recall and production of complex structures.

The absence of notable interaction between explicit instruction through collaborative tasks and the types of feedback regarding learners' ability to recall and produce complex structures can be attributed to the essential function of corrective feedback in collaborative tasks. As a result, these tasks do not create a significant distinction between groups that receive feedback from teachers and those that receive feedback from peers. In essence, because corrective feedback is inherently included in collaborative tasks, these tasks function comparably for both teacher and peer feedback groups. In interpreting these results, it is also worth noting that since corrective feedback types and collaborative tasks act through common cognitive mechanisms, corrective feedback types do not mediate the effect of collaborative tasks on recall and production of complex structures.

In comparing these results with previous studies, since this study was a pioneering investigation into the interaction between explicit instruction of collaborative tasks and types of feedback on learners' recall and production of complex structures, the researcher was unable to identify a comparable study within the current literature on this topic.

6. Conclusion

According to the results, it is concluded that if collaborative tasks (co-practice task writer, corrective feedback providers, and evaluators) are explicitly used, EFL learners' recall and production of complex structures are significantly improved. Moreover, it is concluded that using corrective feedback providers leads to the highest positive effect on EFL learners' recall and production of complex structures. The findings are sufficiently compelling to assert that different forms of feedback, specifically teacher feedback compared to peer feedback, could be fruitful. Accordingly, it can be inferred that feedback provided by teachers is more beneficial than that given by peers in aiding EFL learners in recalling and producing these intricate structures.

The findings hold considerable importance for a range of stakeholders, including EFL educators, students, educational policymakers, and researchers. EFL teachers can utilize collaborative tasks and various types of corrective feedback when teaching complex structures to facilitate the production and recall of these structures for students. Future directions may be explored to benefit from the results of this study and reduce the challenges learners face when using collaborative tasks in complex structures. They can also devise new, effective strategies to cope with the challenges of using collaborative tasks in teaching complex structures and conducting further investigations.

Authords' Contribution

Mozhgan Ebrahimain, the PhD candidate, prepared the literature review and data collection. Bahman Gorjian, the supervisor, helped the team to write the Introduction, methodology, and results of the study. Zohreh Seifoori, the advisor, dealt with the discussion and conclusion sections.

References

Abdolmanafi Rokni, S. J., & Seifi, A. (2014). Dialog journal writing and its effect on learners' speaking accuracy and Fluency. *Study in English Language Teaching, 2*(1), 28-37. http://doi.org/10.22158/selt.v2n1p28

Braund, H., & DeLuca, C. (2018). Elementary students as active agents in their learning: an empirical study of the connections between assessment practices and student metacognition. *The Australian Educational Researcher*, *45*(1), 65-85. http://doi.org/10.1007/s13384-018-0265-z

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Creswell, J. W. (2020). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research, global edition*. Pearson.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal: An Electronic Refereed Journal for Foreign and Second Language Educators, 1*(1), 3-18. http://doi.org/10.5070/l2.v1i1.9054

Farhangi, B., & Pourmihammadi, M. (2018). The effect of intensive teaching of conditional sentence type one on Iranian grade three high school students' writing ability. *International Journal of English Language & Translation Studies, 6*(3), 99-105.

Ghoorchaei, B., Mamashloo, F., Ayatollahi, M. A., & Mohammadzadeh, A. (2022). Effect of direct and indirect corrective feedback on Iranian EFL writers' short and long-term retention of subject-verb agreement. *Cogent Education, 9*(1), 1-16. https://doi.org/10.1080/2331186x.2021.2014022

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Izadpanah, J. (2023). The effect of explicit corrective feedback on EFL learners' retention of grammar: Does the medium of feedback matter? *Journal of Studies in Learning and Teaching English Volume, 12*(1), 99-122. https://doi.org/20.1001.1.22518541.2023.12.1.6.4

Javan Amani, Z., Sadeghi, K., & Alavinia, P. (2024). Combined effect of the phase of explicit instruction and feedback timing on task performance and grammar acquisition. *Teaching English Language, 18*(2), 69-92. https://doi.org/10.22132/tel.2024.341440.137

Kang, S. H., McDermott, K. B., & Roediger III, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*(4-5), 528-558. https://doi.org/10.1080/09541440601056620

Kouhsarian, M, Ghoorchaei, B., Mazandarani, O. & Shomooss, N. (2023). How collaborative learning impacts Motivation and vocabulary learning: The case of Iranian high school learners. *Research in English Language Pedagogy, 11*(2), 138-150. https://doi.org/10.30486/RELP.2023.1979883.1438

Kourang Beheshti, E., & Sadeghi, E. (2019). The effect of using mobile apps on the acquisition of conditional sentences among Iranian intermediate EFL learners. *Iranian Journal of English for Academic Purposes, 8*(1), 1-16. https://doi.org/ 20.1001.1.24763187.2019.8.1.1.1

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309–365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Lightbown, P. M., & Spada, N. (2020). Teaching and learning L2 in the classroom: It's about time. *Language Teaching*, *53*(4), 422-432. https://doi.org/10.1017/S0261444819000454

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 20*(1), 37-66. http://doi.org/10.1017/S0272263197001034

Marefat, H., & Abdollahnejad, E. (2014). Acquisition of English relative clauses by adult Persian learners: Focus on presumptive pronouns. *Journal of Teaching Language Skills, 5*(4),19-40. https://doi.org/10.22099/JTLS.2014.1858

Pakseresht Mogharab, M., Khodabandehlou, M., & Jahandar, Sh. (2014). A comparative study of teaching conditionals at the high school level: The effect of explicit vs. implicit teaching on Iranian EFL

learners' knowledge of conditional sentences. *Indian Journal of Fundamental and Applied Life Sciences, 4*(1), 371-377.

Rezai, M. J. (2011). Acquisition of resumptive pronouns in interrogative embedded clauses by Persian learners of English. *Iranian EFL Journal, 7*(4), 24-44.

Rezapoor, F., & Mohammadzadeh, A. (2024). Investigating the relationship between teacher resilience, teacher immunity and emotional intelligence among Iranian EFL teachers. *Research in English Language Pedagogy, 12*(1), 77-97. https://doi.org/10.30486/RELP.2023.1989926.1479

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris, & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.13.09val

Seifoori, Z., & Fattahi, J. (2014). The comparison of the method section of applied linguistics articles written by native and Iranian writers in terms of grammatical complexity and clause types. *Procedia-Social and Behavioral Sciences, 98*, 1698–1705. https://doi.org/10.1016/j.sbspro.2014.03.596

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*(1), 263-300.https://doi.org/10.1191/1362168804lr146oa

Sippel, L. (2021). Maximizing the benefits of peer interaction: Form-focused instruction and peer feedback training. *Language Teaching Research, 8*, Article 136216882110046. https://doi.org/10.1177/13621688211004638

Van den Berg, B. A. M., Van der Pol, J., Admiraal, W. F., & Simons, P. R. J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers & Education*, *51*(4), 1804-1817. https://doi.org/10.1016/j.compedu.2008.06.001

Wang, C. (2024). Six-minute flipped classroom: The effect of using podcasts and collaborative learning on Taiwanese freshmen's English communicative competence. *Teaching English Language, 18*(1), 1-23. https://doi.org/10.22132/tel.2024.444021.15

Wistner, B., Hideki, S., & Mariko, A. (2013). *An analysis of the Oxford Placement Test and the Michigan English Placement Test as L2 proficiency tests*. Retrieved from https://www.hosei.ac.jp/bungaku/museum/html/kiyo/58/articles/Wistner.pdf