

CHAPTER FOUR
ANALYSES AND RESULTS

The present study attempts to seek answers to several questions concerning the effectiveness of two different types of written CF (direct vs. indirect), which were raised in the first chapter of this thesis. In this chapter, the researcher presents the results of the quantitative analyses of the data. For this purpose, first, the descriptive statistics for the dependent variables in the pretests are tabulated and described. In the next step, the findings which were obtained from the quantitative analyses of the immediate posttest and delayed posttest data based on inferential statistics are presented.

4.1. Pretest Results

4.1.1. Proficiency Test

The results of the Nelson Proficiency Test for elementary levels which was given to the learners in the experimental and control groups prior to the treatment are presented in Table 4.1. This proficiency test was administered before the intervention to make sure that the 3 groups in the study were homogeneous as far as language proficiency is concerned.

Table 4.1
Descriptive statistics for the proficiency test

Groups	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Direct	20	34.45	8.611	1.926	20	48
Indirect	20	38.40	8.152	1.823	19	47
Control	20	32.90	6.248	1.397	20	44
Total	60	35.25	7.957	1.027	19	48

As the descriptive data in Table 4.1 shows, different groups' mean scores on general linguistic proficiency ranged from 32 to 38 out of 50. The analyses revealed that indirect group performed better than other groups with regard to obtained mean score and the control group gained the lowest mean score. Additionally, as the

standard deviations indicate, the control group had the lowest within-group variability but the direct and indirect groups had the highest. In order to find out whether there was a statistically significant difference between the groups, a one-way analysis of variance (ANOVA) was run on the results of the proficiency test. The results indicated that with regard to general linguistic proficiency, there was not a statistically significant difference between the direct group ($M = 34.45$, $SD = 8.61$), indirect group ($M = 38.40$, $SD = 8.15$) and the control group ($M = 32.90$, $SD = 6.24$) at the $p < .05$ level, $F(2, 57) = 2.68$, $p = .07$. Therefore, based on the results of the one-way ANOVA, it was found that the three groups were homogeneous in terms of language proficiency at the outset of the study. The results of the analysis of variance are presented in Table 4.2 below.

Table 4.2
One-way ANOVA for the proficiency test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	321.700	2	160.850	2.686	.077
Within Groups	3413.550	57	59.887		
Total	3735.250	59			

4.1.2. Untimed Grammaticality Judgment Test

Table 4.3 provides the descriptive statistics for the pretest scores from the untimed grammaticality judgment task (UGJT). As can be seen in the Table 4.3, on the pretest for UGJT, the groups averaged 9.40 to 11 out of 30. The mean score for direct correction group (11) and indirect group (10.05) were higher than the mean score for the control group (9.40).

With regard to the standard deviations, all the groups were almost within the same range. That is, for direct, indirect and control group the standard deviation is (2.67), (2.98), and (2.03) respectively. Thus, the control group has the lowest variability and indirect group has the highest within-group variability.

Table 4.3
Descriptive statistics for pretest scores on UGJT

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Direct	20	11.00	2.675	.598	7	16
Indirect	20	10.05	2.982	.667	5	16
Control	20	9.40	2.037	.455	6	14
Total	60	10.15	2.635	.340	5	16

In order to decide whether there was a statistically significant difference between the mean scores of different groups in pretest, a one-way ANOVA was conducted on the mean scores of different groups. The results are displayed in Table 4.4.

Table 4.4
The results of One-way ANOVA for UGJT mean scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25.900	2	12.950	1.924	.155
Within Groups	383.750	57	6.732		
Total	409.650	59			

As shown in Table 4.4, the differences in mean scores of UGJT between the direct group ($M = 11.00$, $SD = 2.67$), indirect group ($M = 10.05$, $SD = 2.98$) and the control group ($M = 9.40$, $SD = 2.03$) were not statistically significant at the $p < .05$ level, $F(2, 57) = 1.92$, $P = .15$. Therefore, with regard to the untimed grammaticality judgment test, the statistical analysis of pretest mean scores of the groups revealed that the three groups were almost homogeneous in terms of their pretest scores at the outset of the study.

4.1.3. Essay Test (Articles)

With regard to the target language use of articles in Learners' essays (story writing), the results of the pretests for essays ranged from 63.70 to 70.05. The results are displayed in Table 4.5 below. As can be seen in the Table 4.5, the highest mean scores

belongs to the indirect group (70.05) whereas the direct group and the control have lower mean scores (68.40 and 63.70). The groups' standard deviations reveal that there was high within-group variability. Control group has the highest standard deviation (23.78) and indirect group has the lowest (18.62).

Table 4.5
Descriptive statistics for pretest scores on the use of articles Essay Task

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Direct	20	68.40	20.620	4.611	18	94
Indirect	20	70.05	18.628	4.165	25	95
Control	20	63.70	23.780	5.317	25	95
Total	60	67.38	20.932	2.702	18	95

The results of a one-way ANOVA which was conducted on the pretest mean scores on learners' essays revealed that differences between the mean scores of direct group ($M = 68.40$, $SD = 20.62$), indirect group ($M = 70.05$, $SD = 18.62$), and control group ($M = 63.70$, $SD = 23.78$) were not statistically significant, $F(2, 57) = .487$, $P = .61$. The results of the ANOVA test are displayed in Table 4.6 below.

Table 4.6
One-way ANOVA for pretest mean scores on the use of articles in essay test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	434.233	2	217.117	.487	.617
Within Groups	25415.950	57	445.894		
Total	25850.183	59			

4.1.4. Essay Test (Simple Past Tense)

Table 4.7 provides the descriptive statistics for pretest mean scores that emerged on the target language use of simple past tense in learners' essays. As can be seen in the Table 4.7, on the pretest the target language use of simple past tense the groups averaged between 57.05 to 63.85. Direct correction group had the lowest mean score (57.05) and the indirect group (63.85) had the highest mean. The mean score for the control group was 61.30.

Table 4.7
Descriptive statistics for pretest scores on the use of simple past tense in essay test

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Direct	20	57.05	27.271	6.098	6	91
Indirect	20	63.85	24.799	5.545	15	94
Control	20	61.30	19.507	4.362	26	93
Total	60	60.73	23.835	3.077	6	94

In order to find whether the groups were homogenous or not in term of their correct use of simple past tense in pretest scores on essay writing, a one-way ANOVA was run on the pretest mean scores of the experimental groups and the control group. The results displayed in Table 4.8, indicated that differences between the mean scores of direct group ($M = 57.05$, $SD = 27.27$), indirect group ($M = 63.85$, $SD = 24.79$), and control group ($M = 61.30$, $SD = 19.50$) were not statistically significant at the $p < .05$ level, $F(2, 57) = .407$, $P = .66$.

Table 4.8
One-way ANOVA for pretest scores on the use of simple past tense in essay test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	472.033	2	236.017	.407	.667
Within Groups	33045.700	57	579.749		
Total	33517.733	59			

4.2. Posttest Results

4.2.1. Untimed Grammaticality Judgment Test

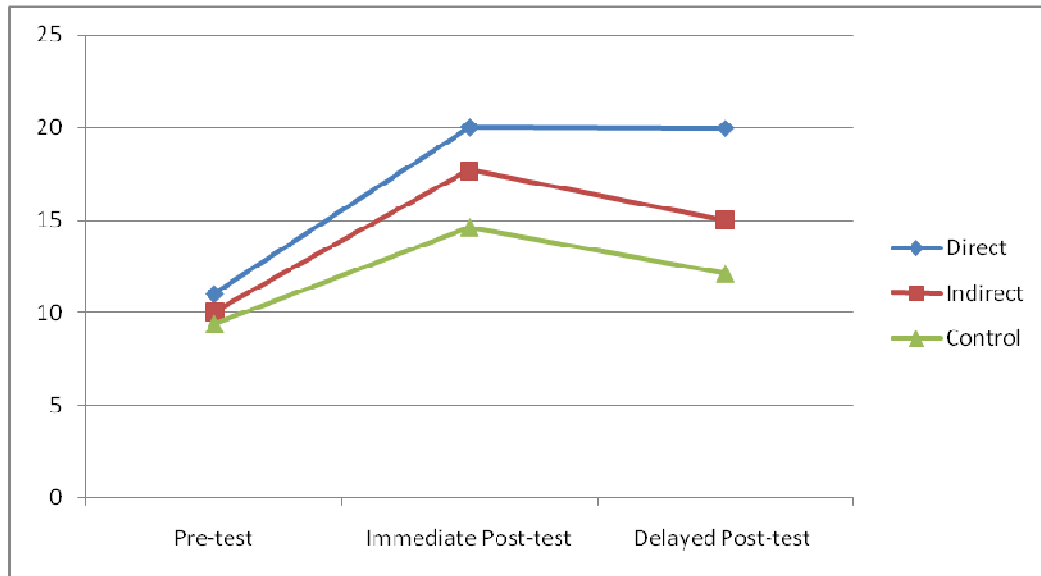
Table 4.9 below compares the mean scores and standard deviations of mean scores in pretest and posttest on the use of articles and simple past tense for Untimed Grammaticality Judgment Test.

Table 4.9
Descriptive statistics for pretest and posttest mean scores on UGJT

CF groups		Pretest	Immediate Post-test	Delayed Post-test
Direct	N	20	20	20
	Mean	11.00	20.00	19.95
	Std. Deviation	2.675	3.372	3.591
Indirect	N	20	20	20
	Mean	10.05	17.70	15.00
	Std. Deviation	2.982	3.881	2.224
Control	N	20	20	20
	Mean	9.40	14.60	12.15
	Std. Deviation	2.037	2.280	3.392
Total	N	60	60	60
	Mean	10.15	17.43	15.70
	Std. Deviation	2.635	3.894	4.473

As can be seen Table 4.9, overall, the learners' knowledge of articles and simple past tense increased from pretest to posttest. Figure 4.1 presents a visual representation of learners' performance on UGJT from pretest to posttest.

Figure 4.1
Groups' mean scores for UGJT across three testing times



A quick glimpse at Figure 4.1 reveals that all the groups improved from pretest to posttest. However, the direct group showed a sharp improvement across the three testing times. Close to the direct group, a rise in the knowledge of articles and simple past tense can be seen in the indirect group and the control group. However, looking at the delayed posttest results, contrary to direct group, we can see a slight drop from immediate posttest to delayed posttest in indirect and control group. In order to examine the differential effects of the different types of written CF treatments, a two way repeated measures ANOVA was run with pretest scores as the dependent variable, and with written corrective feedback treatment (three levels) as independent variables. The results are displayed in table 4.10 and 4.11

Table 4.10**Repeated measures ANOVA for UGJT mean scores across three testing times**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	924.075	1	924.075	133.745	.000
	813.003	1	813.003	86.780	.000
Time * CF	197.600	2	98.800	14.300	.000
	12.156	2	6.078	.649	.527
Error(Time)	393.825	57	6.909		
	534.008	57	9.369		

Table 4.11**Repeated measures ANOVA for UGJT mean scores across three groups**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	37468.939	1	37468.939	3476.028	.000
CF	732.978	2	366.489	34.000	.000
Error	614.417	57	10.779		

As can be seen in Table 4.10, there was a main effect for time and also a statistically significant interaction between time and group. This indicates that the different groups in the study underwent change differently across different testing times. Table 4.11 shows a significant main effect for time. One-way ANOVA with *post hoc* comparisons were conducted on the groups' mean gain scores from pretests to immediate posttest to find out where the differences in groups lied. The results are presented in table 4.12 and 4.13

Table 4.12
One-way ANOVA for gain scores from pretest to immediate posttest (UGJT)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	148.433	2	74.217	4.389	.017
Within Groups	963.750	57	16.908		
Total	1112.183	59			

As can be seen from the Table above, the results of the one-way ANOVA on the mean gain scores of the groups in the immediate posttest of UGJT indicate that there were statistically significant group differences at $p < .05$. $F(2, 57) = 4.38, P < .05$. Post hoc analyses using Tukey's *post hoc* test. Table 4.13 below indicates that in the immediate posttest, the direct group outperformed the control group. However, the indirect group was not significantly different from the control group and the direct group.

Table 4.13
Post hoc Tukey test for immediate posttest scores on UGJT

CF	N	Subset for alpha = 0.05	
		1	2
Control	20	5.2000	
Indirect	20	7.6500	7.6500
Direct	20		9.0000
Sig.		.153	.556

The results of the analysis of variance on the delayed posttest of UGJT which is displayed in Table 4.14 indicated that there was a statistically significant difference between the groups in the post test, $F(2, 57) = 14.30, P < .05$.

Table 4.14**One-way ANOVA for gain scores from pretest to delayed posttest (UGJT)**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	395.200	2	197.600	14.300	.000
Within Groups	787.650	57	13.818		
Total	1182.850	59			

Tukey *post hoc* comparison revealed that in the delayed posttest, the direct group was superior to all the other groups.

Table 4.15**Post hoc Tukey test for posttest gain scores on UGJT**

CF	N	Subset for alpha = 0.05	
		1	2
Control	20	2.7500	
Indirect	20	4.9500	
Direct	20		8.9500
Sig.		.156	1.000

In summary, the results of the repeated-measures ANOVA and subsequent one way ANOVAS on the group mean gain scores on UGJT over different testing times (time 1, time 2, time 3) indicated that there was improvement across the testing times in groups' knowledge of articles and simple past tense and the groups performed differently at different posttests. While in the immediate posttest, both the direct and indirect group performed equally well, and the direct group outperformed the control group, the delayed posttest favored direct CF more than other groups.

4.2.2. Essay Test (Articles)

Table 4.16

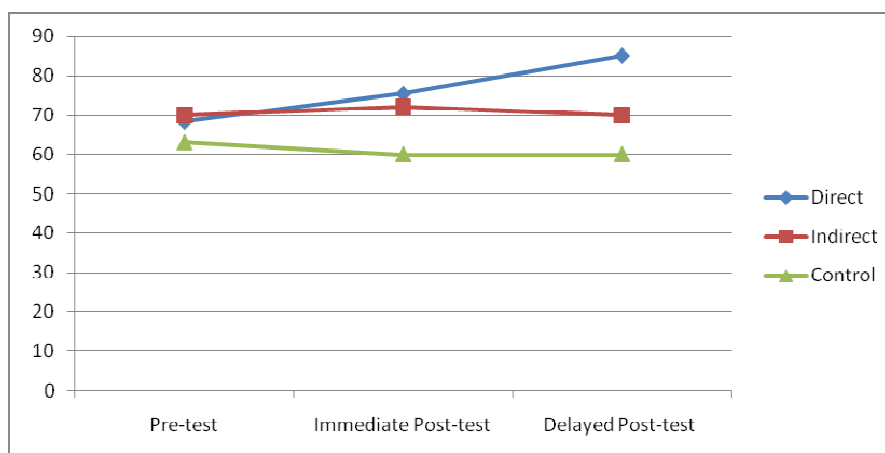
Descriptive statistics for pretest and posttest scores on the use of articles in essays

Written CF groups		Pretest	Immediate Posttest	Delayed posttest
Direct	N	20	20	20
	Mean	68.40	75.40	85.00
	Std. Deviation	20.620	15.267	17.941
Indirect	N	20	20	20
	Mean	70.05	72.45	70.50
	Std. Deviation	18.628	18.619	23.141
Control	N	20	20	20
	Mean	63.70	60.30	60.90
	Std. Deviation	23.780	15.848	24.662
Total	N	60	60	60
	Mean	67.38	69.38	72.13
	Std. Deviation	20.932	17.636	23.912

It can be seen from the data in Table 4.16 that, overall, the learners' knowledge of articles and simple past tense increased from pretest to posttest in direct group. Figure 4.2 presents a visual representation of learners' performance on UGJT from pretest to posttest.

Figure 4.2

Groups' mean scores for the use of articles in essays across three testing times



As Figure 4.2 shows, the mean scores for the direct correction group had a slight rise from pretest to immediate posttest and a tangible change from immediate posttest to delayed posttest. However, the other groups did not show a considerable improvement. Control group had a slight decrease from pretest to the delayed posttest. A repeated measure ANOVA was performed on the pretest and posttest mean scores. The results shown in Tables 4.17 and 4.18 indicated that there was not a significant time effect $F(1, 57) = 2.34, P > .05$, but a significant group effect $F(2, 57) = 1607.54, P < .005$, as well as an interaction effect for both time and group, $F(2, 57) = 3.74, P < .05$.

Table 4.17

Repeated measures ANOVA for the use of articles in essays mean scores across three testing times

Source	TIME	Type III Sum of Squares	df	Mean Square	F	Sig.
TIME	Linear	676.875	1	676.875	2.348	.131
	Quadratic	5.625	1	5.625	.015	.904
TIME * CF	Linear	2159.150	2	1079.575	3.745	.030
	Quadratic	133.317	2	66.658	.175	.840
Error (TIME)	Linear	16430.475	57	288.254		
	Quadratic	21673.225	57	380.232		

Table 4.18

Repeated measures ANOVA for the use of articles in essays across the three groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	872784.200	1	872784.200	1607.545	.000
CF	6592.133	2	3296.067	6.071	.004
Error	30947.000	57	542.930		

The gain scores for different groups from pretest to immediate posttests were submitted to one-way ANOVA to find which group or groups performed differently from others from pretest to posttest. Table 4.19 shows the results.

Table 4.19**One-way ANOVA for gain scores from pretest to immediate posttest (Essay articles)**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1086.400	2	543.200	.800	.454
Within Groups	38699.600	57	678.940		
Total	39786.000	59			

The results of one-way ANOVA on the mean gain scores from pretest to posttest revealed that there was not a statistically significant difference between the three groups in the study, $F(2,57) = .800, P > .05$. A One-way ANOVA was also run on the groups' mean gain scores from pretest to delayed posttest. Table 4.20 shows the results.

Table 4.20**One-way ANOVA for gain scores from pretest to delayed posttest (Essay articles)**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4318.300	2	2159.150	3.745	.030
Within Groups	32860.950	57	576.508		
Total	37179.250	59			

The results of one-way ANOVA of the groups' gain scores revealed that there was a significant difference between the groups' gain scores from pretest to delayed posttest, $F(2,57) = 3.74, P < .05$. The *post hoc* Tukey test (see Table 4.21) indicated that the direct group differed significantly from the other groups.

Table 4.21**Post hoc Tukey test for gain scores from pretest to delayed posttest (Essay articles)**

CF	N	Subset for alpha = 0.05	
		1	2
Control	20	-2.8000	
Indirect	20	.4500	.4500
Direct	20		16.6000
Sig.		.904	.093

4.2.3. Essay Test (Simple Past Tense)

Table 4.22 provides the mean scores and standard deviations for the use of past tense in essays in the immediate posttests and delayed posttests in comparison with the pretest scores.

Table 4.22
Descriptive statistics for pretest and posttest scores on the use of past tense in essays

Written CF groups		Pretest	Immediate Posttest	Delayed posttest
Direct	N	20	20	20
	Mean	57.00	71.70	83.10
	Std. Deviation	25.735	26.569	19.221
Indirect	N	20	20	20
	Mean	63.85	83.05	76.55
	Std. Deviation	24.799	12.580	14.717
Control	N	20	20	20
	Mean	61.30	61.25	62.45
	Std. Deviation	19.507	30.215	19.411
Total	N	60	60	60
	Mean	60.72	72.00	74.03
	Std. Deviation	23.281	25.552	19.636

As it is apparent in Table 4.22, overall, the learners' ability to use simple past tense underwent some change from pretest to immediate posttest in the experimental groups. But the control group's mean scores did not have a considerable change. Figure 4.3 provides a visual representation of different groups' performance from pretests to the immediate posttest and delayed posttests.

As can be seen in Figure 4.3, the indirect groups improved from pretest to the immediate posttest. However, from immediate posttest to delayed posttest, we had a decline in performance. Similar to the indirect group, the direct group underwent an improvement from pretest to posttest but contrary to indirect group, this improvement continued into the delayed posttest. Looking at the pattern of improvement in the control group, we can see no pattern of improvement. A two-way repeated measures ANOVA was conducted on the pretest and posttest scores. The results are shown in Tables 4.23 and 4.24.

Figure 4.3

Groups' mean scores for the use of articles in essays across three testing times

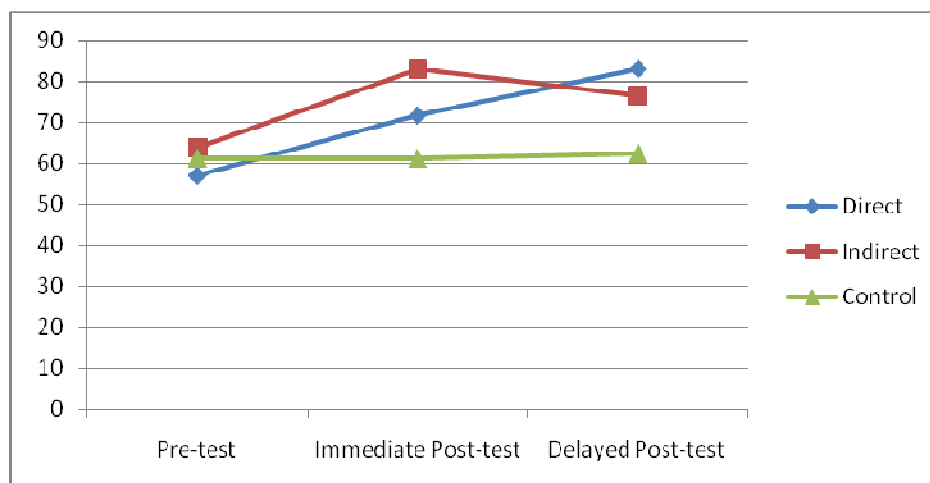


Table 4.23

Repeated measures ANOVA for the use of simple past tense in essays mean scores across the three testing times

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	Linear	5320.008	1	5320.008	29.946	.000
	Quadratic	855.625	1	855.625	2.746	.103
Time * CF	Linear	3118.217	2	1559.108	8.776	.000
	Quadratic	1387.517	2	693.758	2.227	.117
Error(Time)	Linear	10126.275	57	177.654		
	Quadratic	17757.692	57	311.538		

Table 4.24

Repeated measures ANOVA for the use of articles in essays across the three groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	854911.250	1	854911.250	875.283	.000
CF	5183.033	2	2591.517	2.653	.079
Error	55673.383	57	976.726		

The results of the two-way repeated measures ANOVA revealed that overall there was a statistically significant increase from pretest to the posttests, while the interaction effect shows that the increase varied among groups. In order to find which group or groups improved significantly across the testing times, the gain scores for different groups from pretest to posttest were submitted to one way ANOVA. The results are displayed in Table 4.25.

Table 4.25
One-way ANOVA for gain scores from pretest to immediate posttest (Essay articles)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4055.833	2	2027.917	3.513	.036
Within Groups	32904.350	57	577.269		
Total	36960.183	59			

The results of ANOVA analysis, revealed that there was a statistically significant difference between the groups' mean gain scores in the use of simple past tense from pretest to posttest, $F(2, 57) = 3.51, P < .005$. *Post hoc* Tukey comparisons test (see Table 4.26) indicated that there was a statistically significant difference between indirect group and the control group. However, the direct group's difference from the control group and indirect group was not statistically significant.

Table 4.26
Post hoc Tukey test for gain scores from pretest to immediate posttest (Past tense)

CF	N	Subset for alpha = 0.05	
		1	2
Control	20	-.0500	
Direct	20	14.7000	14.7000
Indirect	20		19.2000
Sig.		.137	.825

A one- way ANOVA was also run on the mean gain scores from pretests to posttests. The results are displayed in Table 4.27

Table 4.27**One-way ANOVA for gain scores from pretest to delayed posttest (Essay articles)**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6236.433	2	3118.217	8.776	.000
Within Groups	20252.550	57	355.308		
Total	26488.983	59			

As is displayed in Table 4.27, the ANOVA results revealed a statistically significant difference between the groups' mean gain scores in the use of simple past tense from pretest to delayed posttest, significant, $F(2, 57) = 8.77, P < .005$. *Post hoc* Tukey comparisons test (see Table 4.28) indicated that there was a statistically significant difference between direct group and the control group. However, the indirect group's difference from the control group and direct group was not statistically significant.

Table 4.28**Post hoc Tukey test for gain scores from pretest to delayed posttest (Past tense)**

CF	N	Subset for alpha = 0.05	
		1	2
Control	20	1.1500	
Indirect	20	12.7000	12.7000
Direct	20		26.1000
Sig.		.137	.072

In summary, the results of the statistical tests conducted on the data showed that in the two measuring instruments of this study, the experimental groups and the control group performed differently from each other on the two linguistic features under study. These revealed that written CF can have differential effects on language learning depending on the type and linguistic features.