The Journal of Applied Linguistics Vol. 2, Issue 1
Spring 2009

A Comparative Study of the Effect of Homogeneous and Heterogeneous Collaborative Interaction on the Development of EFL Learners' Writing Skill

Parviz Maftoon
Nasser Ghafoori[*]
Islamic Azad University Science and Research Branch

This study investigates the effect of homogeneous and heterogeneous peer interaction on the development of Iranian EFL learners' writing skill. Sixty female students of TEFL participated in the study. The participants were divided into two groups based on their English proficiency test scores. The homogeneous group consisted of 14 participants paired with partners with similar English proficiency test scores, while the heterogeneous group consisted of 16 participants who were paired with partners who had higher test scores. The pairs had interaction and peer collaboration before carrying out three types of writing tasks. The Repeated Measures ANOVA was used to compare the student writers' pretest writing scores with their three post-test scores. The results showed that both groups, very similarly, had significantly higher post-test scores in all three writing tasks. The findings are explained based on the sociocultural theory and Vygotsky's notion of the zone of proximal development (ZPD). The study offers several important pedagogical implications and suggestions for further research.

Keywords: Dyadic Interaction, Collaboration, Table Description Task, Free Composition Task, Sociocultural Theory

[*] Corresponding Author. Email Add. *: ghafoori_tefl@yahoo.com*

The widespread use of group and pair work in second language (L2) classes is supported by both pedagogic arguments and research findings. Studies have shown that learners working in groups, particularly in cooperative groups, are exposed to a great variety of viewpoints, construct new ways of understanding, and develop greater critical skills (see, for example, Webb, 1989).

Psycholinguistic theories of L2 acquisition , in particular research findings based on Long's (1983, 1996) Interaction Hypothesis and Krashen's (1982, 1985) Input Hypothesis, have provided additional support for the use of group and pair work in L2 classrooms. Long argues that interaction promotes acquisition because interaction provides learners with the comprehensible input needed for acquisition to take place. Long (1996) emphasizes the importance of negative feedback and modified output for L2 learning. Negative feedback can be explicit (e.g., explicit correction) or implicit (e.g., clarification requests, recasts, etc.), serving a role in raising learners' awareness of the problematic aspects of their utterances.

Research guided by this theoretical perspective (e.g., Pica, Young & Doughty, 1987) has shown that, if careful attention is paid to the tasks used, and the strategic grouping of students in terms of gender, familiarity, and L2 proficiency, small group work provides learners with opportunities to give and receive feedback. Thus, from this theoretical perspective, it is interaction between learners that helps L2 learning. According to Pica (1994), the process of "the modification and restructuring of interaction that occurs when learners and their interlocutors anticipate, perceive, or experience difficulties in message comprehensibility" (p.495) has been referred to as negotiation". Through negotiation, comprehensibility is achieved as interlocutors repeat and rephrase for their conversational partners.

Lightbown and Spada (1999), Doughty and Williams (1998), among others, have explored how interaction provides opportunities for learners not only to negotiate the message of the input, but, in so doing, to focus on its form as well. Other

researchers, for example, Aljaafreh and Lantolf (1994), and Nassaji and Swain (2000) have examined the nature and type of feedback that may be most helpful to learners during interaction at different stages of their acquisition of language form.

Swain (2000) has proposed another hypothesis emphasizing the role of output and production in L2 learning. She argues that "output pushes learners to process language more deeply – with more mental effort – than does input" (p. 99). Her point is that in speaking or writing, learners need to do something. They need to create linguistic form and meaning, and in so doing, discover what they can and cannot do. According to this hypothesis, learners' meaningful production of language – output – seems to have a potentially significant role in language development. Swain (2000) has also introduced the notion of collaborative dialogue "in which speakers are engaged in problem solving and knowledge building" (p. 102). She shows through examples that collaborative dialogue mediates joint problem solving and knowledge building. She writes:

> When a collaborative effort is being made by participants in an activity, their speaking (or writing) mediates this effort. As each participant speaks, their 'saying' becomes 'what they said', providing an object for reflection. Their 'saying' is a cognitive activity, and 'what is said' is an outcome of that activity. Through saying and reflecting on what was said, new knowledge is constructed. (Swain, 2000, p. 113)

However, the negotiation research has recently taken a new approach to the study of interaction under the theoretical framework referred to as socio-cultural theory, which is based on the work of Vygotsky (1978, 1981, 1986) and neo-Vygotskian scholars such as Leontiev (1981) and Wertsch (1985, 1991).

Researchers advocating sociocultural framework (e.g., Lantolf, 2000, 2006, 2007; Lantolf & Thorne, 2006) call for research, which is conducted in the context in which language is actually taught, namely the L2 classroom. In the sociocultural

school of thought, language is seen as a psychological tool which mediates learning and social relations. According to Lantolf (2000) "the most fundamental concept of sociocultural theory is that the human mind is mediated" (p. 1). It means that just as humans, we do not act directly on physical world but rely, instead, on tools and labor activity, which allows us to change the world, and with it, the circumstances under which we live in. We also use symbolic tools, or signs, to mediate and regulate our relationships with others and with ourselves, and thus change the nature of these relationships. Physical as well as symbolic tools are artifacts created by human cultures over time and made available to succeeding generations, which can modify these artifacts before passing them on to future generations. Included among symbolic tools are numbers and arithmetic systems, music, art, and above all language. (Lantolf, 2000).

According to Vygotsky (1978, 1981), human cognitive development is inherently a socially situated activity. A child's (novice) cognitive development arises in social interactions with a more able member of society (e.g., parent, peer). The more able member (expert), by providing the novice with the appropriate level of assistance, often referred to as "scaffolding", enables the novice to reach a higher level of development. The difference between the novice's actual level of development and the potential level reached with the assistance of the expert is termed as the Zone of Proximal Development (ZPD). In Vygotsky's (1978) definition, ZPD is "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance in collaboration with more capable peers" (p. 86).

Although Vygotsky's work focused on the cognitive development of children, it is argued that the theory is applicable to all kinds of learning, child and adult, within formal and informal instructional settings, and in asymmetrical (expert-novice) as well as symmetrical (equal ability) groupings (van Lier, 2000). In the field of L2 research, studies have also shown that scaffolding and assistance can occur not only in teacher-learner interaction (e.g.,

Aljaafreh & Lantolf, 1994) but also in peer interaction, when learners work in small groups or pairs (e.g., Donate 1994; Nassaji & Swain, 2000)

Studies comparing individual and pair work on writing and grammar-focused tasks have shown some advantage for pair-work. Storch (1999) compared individual and pair performance on a range of grammar-focused exercises (multiple-choice, cloze, text reconstruction) and found that pairs completed the exercises more accurately than the students who worked alone. However, since the same students performed the exercises in pairs and individually, the results might have been confounded by a practice effect.

In another study, Storch (2005) compared short reports produced by students working in pairs with those working individually, and found that pairs produced more grammatically accurate texts. In a more recent study, Storch (2007) investigated the merits of pair work by comparing pair and individual work on a text-editing task. The analysis of the edited texts showed that there were no significant difference between the accuracy of tasks completed individually and those completed in pairs. On the other hand, the analysis of the transcribed pair talk showed that most pairs engaged actively in deliberations over language and tended to reach correct resolutions. Thus, the results suggested that, "although pair work on a grammar-focused task may not lead to greater accuracy in completing the task, pair work provided learners with more opportunities to use the second language for a range of functions, and in turn for language learning" (p. 143).

Given the small body of existing research on the efficiency of small group and pair work to enhance L2 development, there is clearly a need for further research in the field. Such research is also needed because not all language learners (nor teachers according to McDonough, 2004) seem convinced of the merits of small group and pair work. As a language teacher, we have often observed when we ask students to work in pairs or in small groups, particularly on tasks that require written output and are more grammar focused, some students seem reluctant to work in pairs, preferring to complete the tasks individually. This may be due to the fact that working in groups or pairs is a culture-bound

phenomenon.

A number of survey studies support these observations. In a study of ESL high school students' preferences conducted in Australia, Mishra and Oliver (1998) found that although 70% of the students had positive attitudes toward group and pair work, very few learners, especially from South-East Asia, liked group and pair work on grammar-focused tasks. The students preferred to work individually on such tasks because they felt that this could provide them with more opportunities to practice their grammar. Kinsella (1996) reported on ESL students' concerns about learning the "wrong" grammar from their peers. Similar concerns were voiced by Thai EFL learners in interviews conducted by McDonough (2004), particularly when working in small groups on grammar-focused tasks.

Another related issue is the proficiency level of peers in pair and small group works. Almost little research has documented how learners with different language proficiency levels interact with each other, and whether homogenous groups and pairs are more advantageous than heterogeneous ones. Watanabe and Swain (2007) investigated the effects of L2 proficiency differences in pairs and patterns of interaction in L2 learning. They found that proficiency differences do not necessarily affect the nature of peer assistance and L2 learning.

The purpose of the present study was to investigate the effect of two types of homogeneous and heterogeneous dyadic interaction on the EFL learners' writing ability in three types of writing tasks, Picture Description, Table Description, and Free Composition. The homogeneous pairs were pairs in which the two participants were in similar L2 proficiency level, and the heterogeneous pairs were pairs with two participants who were different in terms of L2 proficiency level. The findings of the study would be of considerable significance in the design of EFL courses in general and writing courses in particular. Pair work and collaborative activities, if proved to be effective in improving the learners' language related skills, can be employed in EFL classes as a powerful and lively means to increase the learners' L2 communicative competence in general and their writing skill in

particular. The study specifically aimed to answer the following questions:

1) Does collaborative interaction between the pairs of the homogeneous group (G-Hom) lead to a difference between the participants' pretest-posttest writing scores?
2) Does collaborative interaction between the pairs in the heterogeneous group (G-Het) lead to a difference between the participants' pretest-posttest writing scores?
3) Is there a difference between G-Hom and G-Het in terms of the pretest-posttest differences?

Method

Participants

Sixty female students who studied Teaching English as a foreign Language (TEFL) in Islamic Azad University, Tabriz Branch took part in this study. For the facility in administration, and in order to keep the gender factor neutral in the study, all the participants were selected from female students. The age range of the participants was 19 to 25. They were all in the fifth, sixth, and seventh semesters of their studies, and they had already passed several basic courses in conversation, grammar, and writing. These participants were selected from the body of 124 students who had taken an English language proficiency test and had been already ranked based on their total test scores. Fourteen students with low English proficiency test scores were paired with partners who had high scores to make the heterogeneous group (G-Het). Also, sixteen students were paired with partners who had similar test scores to make the homogeneous group (G-Hom).

Instrumentation

The instruments used for the data collection in the study included a standardized English language proficiency test and three types of writing tasks/tests.

- English Language Proficiency Test

The First Certificate English (FCE) language proficiency test was used in order to examine the participants' English proficiency at the beginning of the study and to design two experimental groups (G-Hom and G-Het) based on the participants' test scores. The test included five sections of Reading, Writing, Use of English, and Speaking The participants' scores on the writing section of the FCE test were also considered as their writing pretest scores to be compared with their final writing scores in three posttest writing tasks. The Writing Section of the FCE included two parts. Part One, which was a compulsory task, required a transactional letter, i.e. a request for action or response to a request for action, and was based on input materials, such as advertisements or short articles. Part Two was an optional task from a choice of four. The optional tasks were drawn from a variety of genres, including non-transactional letters, discursive compositions, narratives, and descriptions. The examinees were asked to write between 120 and 180 words for each task, and the total time for the test was one hour 30 minutes (Appendix A).

- Writing Tasks

Three types of writing tasks were given to the participants in the study. : Picture Description task in which the participants looked at a picture and talked about the picture in pairs before they wrote a description about it (see Appendix B). , Table Description task in which the participants were required to talk about the information in a given table and then to write a description (see Appendix C), and Composition Task requiring the participants to talk in pairs first about some given topics, and then to write a composition (see Appendix D). Each writing task had four isomorphic versions, which were given to the participants twice a week in four successive weeks. The pairs in both groups were required to carry out the first three versions in each task cycle through peer collaboration and dyadic interaction and the fourth version individually. The same individual in each pair carried out all writings and their scores in the fourth versions of the writings were compared with the student/writers pretest writing scores

obtained from the writing section of the FCE test.

The TEEP (Test in English for Educational Purposes) analytic scale proposed by Weir (1990, cited in Weigle, 2002) was used to score all the writings in the pretest and posttests. The TEEP scheme (see Appendix E) consists of seven scales (A to G), each divided into four levels with score points ranging from 0 to 3. The first four scales are related to communicative effectiveness, and the other three are related to grammatical and mechanical accuracy. The score range of the whole scale is then from 0 to 21.

Design

The research design of the study was quasi-experimental design. The main purpose was to investigate and compare the effect of homogeneous and heterogeneous dyadic interaction on the learners' writing skill. Thus, the independent variable or the treatment in the study was two types of dyadic interaction or peer collaboration, and the dependent variable was the participants' writing ability.

Procedure

The participants' writing scores in the writing section of the FCE test were considered as the pretest scores. The same participants, in two groups, produced 12 writings (four writings for each of the three writing tasks) during the study, and their scores in the fourth version of each writing task, which was carried out individually, were considered as the posttest scores.

In each group, the students/writers' mean scores on the writing section of the FCE test were compared with their mean scores on the fourth writing task in each task cycle. The General Linear Model Repeated Measures ANOVA with LSD Post Hoc Pairwise Comparisons was used to examine the differences between the mean scores of the pretest and three posttests in two groups. All writings were scored by the application of the TEEP scoring scheme.

Eighty samples (almost one third) of the participants'

writings were randomly selected and scored by another rater who was trained by the researcher to work with the TEEP scoring scheme. The second rater's scores were compared with the researcher's scores to examine the inter-rater reliability of the writing scores. As it is shown in Table 1, a correlation coefficient of 0.88 was found as the inter-rater reliability of two scorings.

Table 1
Pearson correlation coefficient showing the inter-rater reliability of the ratings of 80 sample writings

		First	Second
First	Pearson Correlation	1	.878(**)
	Sig. (2-tailed)		.000
	N	80	80
Second	Pearson Correlation	.878(**)	1
	Sig. (2-tailed)	.000	
	N	80	80

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

Furthermore, 50 of the students' writings were randomly selected, and scored for the second time by the researcher in an interval of 30 days to examine the intra-rater reliability of the writing scores. Table 2 shows the Pearson coefficient of correlation between the two ratings. The correlation was found to be 0.92 as the indicator of intra-rater reliability of the writing scores.

Table 2
P*earson correlation coefficient showing the intra-rater reliability of the ratings of 50 sample writing*

		Test	retest
test	Pearson Correlation	1	.926(**)
	Sig. (2-tailed)		.000
	N	50	50
retest	Pearson Correlation	.926(**)	1
	Sig. (2-tailed)	.000	
	N	50	50

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

Results

Three null hypotheses were developed based on the three research questions as follows:

H_0 (1) There is not any statistically significant difference between the participants' pretest-posttest writing scores in G-Hom.

H_0 (2) There is not any statistically significant difference between the participants' pretest-posttest writing scores in G-Het.

H_0 (3) There is not any statistically significant difference between G-Hom and G-Het in terms of their pretest-posttest differences.

The level of significance to reject the null hypotheses was set to be 0.05. The rejection of the null hypotheses would provide answer for the research questions. The Repeated Measures statistics was used to answer the first two research questions, and the Univaiate Analysis of Variance (ANOVA) was used to answer the third research question.

Table 3 shows the G-Hom participants' writing scores in the FCE test (pretest) and the scorers in the final version of three writing tasks of Picture Description (PD), Table Description (TD), and Free Composition (FC).

Table 3
The G-Hom participants' scores in pre-test writing and three post-test writings

No	Participants	Pre-test FCE	Post-test PD	Post-test TD	Post-test FC
1	S. H.	11	14	12	13
2	E. K.	13	16	17	17
3	R. K.	13	17	13	15
4	E. P.	7	13	12	12
5	M. A.	8	11	11	13
6	H. B.	8	9	8	9
7	H. Z.	11	13	15	13
8	S. A.	10	11	11	11
9	N. H.	8	12	13	15
10	E. J.	9	18	16	15
11	N. S.	8	13	15	15
12	N. V.	8	12	12	14
13	T. M.	7	12	11	12
14	E. B.	16	17	17	17
15	M. D.	10	16	16	15
16	M. K.	16	19	18	19
	Mean	10.18	13.93	13.56	14.06

Table 4 shows the results of the Repeated Measures ANOVA tests of within-subjects Effects in G-Hom. The table shows that there are significant differences among the four sets of scores because the P-value observed (F= 31, df, 3) is below the P-value selected for rejection of the null hypotheses.

The null hypothesis (1) could be rejected as there was evidence to show that there was a statistically significant difference among the pretest scores and posttest scores in G-Hom. The LSD Post Hoc test was run to examine the pair wise differences between the G-Hom participants' four sets of scores obtained from the FCE pretest and the three posttests of PD, TD, and FC.

Table 4
Repeated measures test of within-subjects effects in G-Hom

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	Sphericity Assumed	163.500	3	54.500	31.044	.000	.674
	Greenhouse-Geisser	163.500	2.194	74.519	31.044	.000	.674
	Huynh-Feldt	163.500	2.585	63.246	31.044	.000	.674
	Lower-bound	163.500	1.000	163.500	31.044	.000	.674
Error(task)	Sphericity Assumed	79.000	45	1.756			
	Greenhouse-Geisser	79.000	32.911	2.400			
	Huynh-Feldt	79.000	38.777	2.037			
	Lower-bound	79.000	15.000	5.267			

Table 5 shows the results of the Post Hoc test, which examines the significance of pair wise differences between the mean scores obtained from four tasks.

According to the first row of Table 5, the participants' mean scores increased from pretest to all three posttests of PD, TD, and FD. The increase from FCE test to Picture Description test was 3.750,which is statistically significant (P< 0.05). The difference between the FCE pretest and TD posttest mean scores (3.375) was also statistically significant (P<0.05) . Similarly, the difference between the pretest-posttest scores in the Free Composition task (3.875) was statistically significant (P< 0.05).

Table 5
Post hoc pairwise comparisons for G-Hom

(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
FCE	PD	-3.750*	.536	.000	-4.893	-2.607
	TD	-3.375*	.584	.000	-4.619	-2.131
	FC	-3.875*	.539	.000	-5.024	-2.726
PD	FCE	3.750*	.536	.000	2.607	4.893
	TD	.375	.386	.347	-.448	1.198
	FC	-.125	.397	.757	-.970	.720
TD	FCE	3.375*	.584	.000	2.131	4.619
	PD	-.375	.386	.347	-1.198	.448
	FC	-.500	.303	.119	-1.145	.145
FC	FCE	3.875*	.539	.000	2.726	5.024
	PD	.125	.397	.757	-.720	.970
	TD	.500	.303	.119	-.145	1.145

On the other hand, Table 5 shows that the differences between the three posttests were not statistically significant. This finding suggests that pair work affected the participants' scores to a certain extent so that there was no further increase from Picture Description task to the next two posttest tasks though the participants had more interaction between the tasks. This can be because the effect of pair interaction on the learners' language skill is limited to a certain extent beyond which the development may require more time and more practice. This may be a good topic for further research in the field.

Table 6 shows the G-Het participants' scores in the pretest (FCE Writing) and in the fourth version of each writing task cycle (PD, TD, and FC, respectively). The 14 participants in this group, selected from the lowest part of the L2 proficiency test rank order, had been matched with partners who had higher L2 proficiency test scores. Like G-Hom, the participants in G-Het carried out the first three versions of each writing task through collaboration and pair work, but the fourth version individually. The goal was to

investigate whether collaboration with more proficient peers could have any positive effect on the student/writers' writing ability.

Table 6
The G-Het participants' scores in the pretest writing and three posttest writings

No	Participants	Pre-test FCE	Post-test PD	Post-test TD	Post-test FC
1	D. R.	7	13	13	12
2	F. S	6	16	14	15
3	S. H.	8	15	15	17
4	Z. A.	10	16	12	13
5	S. F.	8	15	12	16
6	F. P.	9	16	15	16
7	R. A.	7	14	16	14
8	R. S.	8	13	14	12
9	P. J.	12	16	17	16
10	R. V.	7	13	14	12
11	N. N.	10	15	14	13
12	F. M.	8	14	13	13
13	A. G.	10	15	15	14
14	S. T.	10	13	15	14
	Mean	8.8	14.57	14.21	14.07

Table 7 shows the results of the Repeated Measures tests of within-subjects effects. As it is seen in the first row of Table 7, the differences among the four mean scores obtained from the four writing tasks were statistically significant ($P< 0.05$). The second null hypothesis was also rejected, and it was found that there was a statistically significant difference among the four sets of scores.

Table 7
Repeated measures tests of within-subjects effects in G-Het

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task2	Sphericity Assumed	344.714	3	114.905	75.588	.000	.853
	Greenhouse-Geisser	344.714	2.473	139.387	75.588	.000	.853
	Huynh-Feldt	344.714	3.000	114.905	75.588	.000	.853
	Lower-bound	344.714	1.000	344.714	75.588	.000	.853
Error (task2)	Sphericity Assumed	59.286	39	1.520			
	Greenhouse-Geisser	59.286	32.150	1.844			
	Huynh-Feldt	59.286	39.000	1.520			
	Lower-bound	59.286	13.000	4.560			

LSD Post Hoc statistics was run to examine the significance of the pairwise differences between the four mean scores obtained from the four tasks (FCE, PD, TD, and FC) for G-Het. Table 8 shows the results of the pairwise comparisons.

As it is demonstrated in Table 8 (first row), the differences between the FCE pretest mean scores and all three posttest mean scores in G-Het were statistically significant. The results for G-Het show that the participants' writing scores increased significantly from the pretest of writing to all three posttest writing tasks. The collaborative interaction that the dyads had before they accomplished their writing tasks helped them to improve their writing skill. The positive effect of collaborative dyadic interaction was seen in the student/writers' higher scores in their posttest individual performances.

Table 8

Post hoc pairwise comparisons for G-HET

(I) task2	(J) task	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
FCE	PD	-6.000*	.445	.000	-6.961	-5.039
	TD	-5.643*	.476	.000	-6.671	-4.615
	FC	-5.500*	.562	.000	-6.714	-4.286
PD	FCE	6.000*	.445	.000	5.039	6.961
	TD	.357	.476	.466	-.671	1.385
	FC	.500	.344	.169	-.243	1.243
TD	FCE	5.643*	.476	.000	4.615	6.671
	PD	-.357	.476	.466	-1.385	.671
	FC	.143	.467	.765	-.866	1.152
FC	FCE	5.500*	.562	.000	4.286	6.714
	PD	-.500	.344	.169	-1.243	.243
	TD	-.143	.467	.765	-1.152	.866

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

The similarities and differences between G-Hom and G-Het in terms of the pretest-posttest gains were further investigated in the study. Figure 4.1 shows that the increases in the three posttests of writing in both groups, as compared with the FCE pretest, were almost parallel. The graph also shows how the participants' performances were almost equal in three posttest-writing tasks.

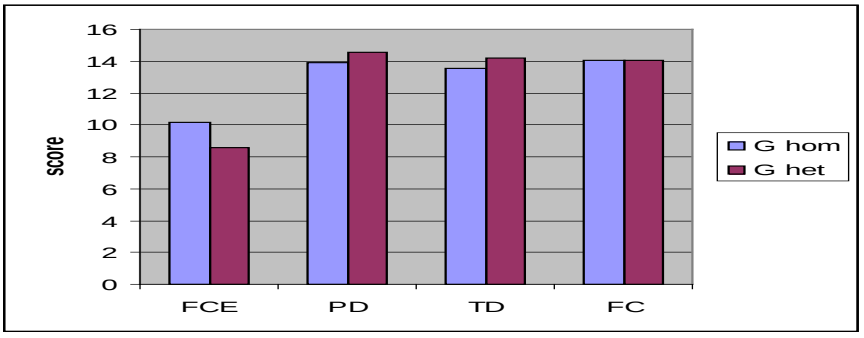


Figure 1. Bar graph comparing the mean scores of G-Hom and G-Het in four tests

Additionally, a Univariate Analysis of Variance (ANOVA) was run to test the third null hypothesis based on the third research question asking whether there was a statistically significant difference between G-Hom and G-Het in terms of their pretest-posttest differences. Table 9 shows the results of the univariate ANOVA tests of between-subjects effects. The rows related to group and group/task interaction show that the F-values observed were smaller than the critical F-values. In other words, the probability level, observed (0.848 and 0.187) for group differences and group/task interaction, respectively, was both greater than the P-value selected (0.05) for rejection of the null hypotheses in the study. Therefore, the null hypothesis (3) stating that there was not any statistically significant difference between the G-Hom and G-Het in terms of the pretest- posttests differences could not be rejected. The answer to the research question (3) was negative. This means there is no statistically significant difference between the two groups in terms of the differences from pretest to posttests.

Table 9
Univariate ANOVA tests of between-subjects effects comparing G-Hom and G-Het pretest-posttest differences

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	508.407[a]	7	72.630	13.920	.000
Intercept	19872.193	1	19872.193	3.809E3	.000
Group	.193	1	.193	.037	.848
Task	494.814	3	164.938	31.611	.000
group * task	25.481	3	8.494	1.628	.187
Error	584.393	112	5.218		
Total	21062.000	120			
Corrected Total	1092.800	119			

a. R Squared = .465 (Adjusted R Squared = .432)

Discussion

The findings about G-Hom show that the pair work the students had before writing their assignments had a positive effect on their writing ability. The collaboration and interaction the participants had with their partners who had almost equal language proficiency level helped them improve their writing skill. As Storch (2007) suggests, "pair work afforded learners opportunities to pool their linguistic resources and co-construct knowledge about language" (p. 155).

The results for G-Het also show that the participants' writing scores increased significantly from the pretest of writing to all three posttest writing tasks. The collaborative interaction that the participants had with partners who were more proficient helped them to improve their writing skill. The positive effect of collaborative dyadic interaction was seen in the student/writers' higher scores in their posttest individual performances.

These findings can be explained in terms of Vygotsky's theory concerning the genesis of higher mental functioning, and, in particular, the process of internalization. Vygotsky's (1978) genetic law of development postulates that the development of

higher mental functions arises out of inter-psychological activity or social activity.

The finding that the differences among the three posttests were not statistically significant both in G-Hom and G-Het suggests that the pair work affected the participants' scores to a certain extent so that there was no further increase from the Picture Description task to the next two posttest tasks though the participants had more interaction between the tasks. This can be because the effect of pair interaction on the learners' language skill is limited to a certain extent beyond which the development may require more time and more practice. This may be a good topic for further research in the field.

The findings are also similar to what other researchers have found regarding the effects of collaboration on the L2 learners' development. Swain et al. (2002), reviewed a number of studies which have shown peer-peer collaborative dialogue to be an important aspect of L2 learning. The researchers believe that collaborative writing and peer revision warrant attention because the collaborative dialogue that emerges in the writing process mediates language learning.

Thus, one explanation for why there were increases in the writing ability of the participants who performed dyadic interaction is that they had "opportunities for learning" (Lantolf, 2000) or "affordances" (van Lier, 2000). In the heterogeneous pairs, the affordance and scaffolding came from the more capable pair, while in the homogeneous pairs the assistance was co-constructed in collaboration but not necessarily with a more capable peer. This is in line with what researchers have already found regarding the benefits of homogeneous pair work (e.g., Donato, 1994; Kowal & Swain 1994; Ohta, 2000, 2001; Swain & Lapkin, 1998; Villamil & de Guerrero, 1996, 1998).

Very similar findings have been reported regarding the benefits of dyadic interaction and peer-peer collaboration both in homogenous and heterogeneous pair work for L2 learning (e.g., de Guerrero & Villamil, 2000; Kim & McDonough, 2008; Leeser, 2004; Storch, 1999, 2001, 2002a, 2002b; Swain & Lapkin, 2002; Watanabe & Swain, 2007).

The similar pattern of improvement from pretest to posttsest in both G-Hom and G-Het shows that although there may be several differences in the patterns of interaction in homogeneous and heterogeneous pairs, the participants in both groups benefited from the dyadic interaction to improve their writing skill. This finding is similar to what Watanbe and Swain (2007) found concerning the effect of L2 proficiency differences on the students' gains. The researchers found that, it was the pattern of interaction, that is, collaborative or dominant/passive, rather than the students' proficiency that was conducive to L2 learning. Overall Watanabe and Swain found that peers of different language proficiency levels could benefit from working with one another, and this finding has also been reported by the previous peer-peer research (e.g., Ohta, 2000; Storch, 2001, Swain & Lapkin, 1998).

The findings provide support for the claim made by sociocultural theory that development, and in this case, language development, is social. However, sociocultural theory tends to focus on heterogeneous pairs (expert-novice), in which partners differ in their level of ability. Yet, the findings of the present study add to the growing body of research (e.g., Donato, 1994; Kim & McDonough, 2008; Storch, 2001, 2002a; Swain, 1998; Watanabe & Swain, 2007) which has shown that learning can also occur in homogeneous peer interaction, that is, interaction between learners who are fairly equal in terms of their level of L2 knowledge. Furthermore, it was found that the type of writing tasks that the participants were required to perform did not seem to play any significant role in their writing performances.

Conclusion

Several important implications can be drawn from the study. Tasks such as writing, which encourage students to reflect on language form while still being oriented to meaning, that is, tasks which, in Swain's (2000) words, engage students in "collaborative dialogue" can be particularly useful for learning strategic processes, as well as lexical/grammatical aspects of the second language. In many of the tasks used in the study of negotiation, the

focus has been on communication where "attention is principally focused on meaning rather than form" (Nunan, 1989, p. 10). However, it is certainly feasible for a communicative task to be one in which learners communicate about language while trying to produce something they want to say in the target language.

Pedagogical Implications

The finding that peers with different as well as similar L2 proficiency levels can benefit from working with one another is a positive finding. It is also in line with what other researchers (e.g., Ohta, 2001; Storch, 2001; Swain & Lapkin, 1998) have found about the benefits of peer-peer interaction and collaboration. It shows that social mediation comes not only from an expert, such as teachers, but also from peers with almost equal ability level, and even from less proficient peers, as Watanabe and Swain (2007) found in their study on adult ESL learners. Therefore, teachers should be careful not to assume that grouping different proficiency peers is less conducive to L2 learning.

In order to facilitate constructive pair work in L2 classrooms, teachers should prepare learners more carefully for pair work. Prior to assigning learners to work in groups or pairs, teachers should engage them in discussions about the advantages of collaboration and model collaborative dialogue. Furthermore, less proficient learners may feel more comfortable interacting with advanced learners if the more advanced learners assume an expert role rather than a dominant role. Therefore, teachers need to encourage advanced learners to act as a facilitator when interacting with peers who are less proficient.

Teachers need to understand the factors that influence students to interact in a certain way and perhaps, more importantly, know how to encourage learners to work collaboratively. A number of suggestions have been made in the literature concerning how teachers can best create "communities of learning" in which collaboration occurs and is valued (see, for example, Lockhart & Ng, 1995). Such suggestions include the need for teachers to emphasize the importance and value of individual learners'

contributions, as well as to present tasks as opportunities for joint learning rather than just as end products for individual assessment.

The present study can also offer some suggestions for further research. An investigation is required to examine the effect of dyadic interaction on the development of other areas of L2 proficiency, such as speaking or reading skills. Moreover, the study examined the question whether more learning occurred when the L2 learners were paired with partners who were more proficient or with partners who were almost equal in terms of L2 proficiency. Whether learners that are more proficient can also benefit from working with less proficient ones is still an open question. Also, the participants' gender was a neutral factor in the present study as all the participants were female. More research is required to examine whether the learners get more benefit when they are matched with partners of the same sex or the opposite sex. In addition, further research can be conducted to investigate whether male or female learners get more benefit from dyadic interaction and peer collaboration.

On the other hand, the writing tasks used in this study to elicit dyadic interaction included Picture Description, Table Description, and Free Composition. As researchers have already shown (e.g., Villamil & de Guerrero, 1998), task differences (for example, expository and persuasive) may influence the performance of the learners during peer revision. Further research is required to examine the effect of collaborative dialogue while doing other types of tasks, such as editing, dictogloss, and error recognition.

Finally, it must be noted that the present study involved a small sample size and was conducted in a controlled classroom and laboratory setting. Before extrapolating from the results of the study to other settings and situations, more research is required with other types of learners and in other learning contexts. It is hoped that this study provides some insight into the complex nature of peer-peer interaction and its importance to L2 learning.

The Authors

Parviz Maftoon is Associate Professor of teaching English at Azad University, Science and Research Branch, Tehran, Iran. He received his Ph.D. degree from New York University in 1978 in Teaching English to Speakers of Other Languages. His primary research interests concern EFL writing, second language acquisition, and syllabus design. He has published and edited a number of research articles and books. He is currently on the editorial board of several language journals in Iran.

Nasser Ghafoori is a lecturer at Islamic Azad University, Tabriz Branch. He received his PhD in Teaching English a Foreign Language (TEFL) from Islamic Azad University, Science and Research Branch. He has presented several papers in a number of local and international conferences. His major research interests are teaching methodology, second language acquisition, assessment, and linguistics.

References

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language leaming in the zone of proximal development. *The Modem Language Journal, 78,* 465-483.

de Guerrero, M. C., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal, 84,* 51-68.

Donato, R. (1994). Collective scaffolding in second language leaming. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33-56). Norwood, NJ: Ablex.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press

Kim, Y. J., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching research, 12,* (2), pp. 211-234.

Kinsella, K. (1996). Designing group work that supports and enhances diverse classroom work style. *TESOL Journal*, 6, 24-30.

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3,* (2), 73-93.

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

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

Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1-26). Oxford: Oxford University Press.

Lantolf, J. P. (2006). Sociocultural theory and L2. State of the art. *Studies in Second Language Acquisition, 28,* 67–109.

Lantolf, J. P. (2007). Sociocultural source of thinking and its relevance for second language acquisition. *Bilingualism: Language and Cognition, 10,* (1), 31-33.

Lantolf, J. P. & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development.* Oxford: Oxford University Press.

Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8,* 55–81.

Leontiev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch, (Ed.), *The concept of activity in Soviet Psychology* (pp. 37-71). Armonk, NY: Sharpe.

Lightbown, P., & Spada, N. (1999). *How languages are learned. (2nd Ed.).* Oxford: Oxford University Press.

Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions and content. *Language learning, 45,* 605-655.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4,* 126-41.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie, & T. Bhatia, (eds.), *Handbook of Second Language Acquisition.* San Diego:

Academic Press.

McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System, 32*, (2), 207-24.

Mishra, S., & Oliver, R. (1998). Secondary school ESL learners' perceptions of pair work in Australian classroom. *TESOL in Context, 8*, 19-23.

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness, 9*, 34-51.

Nunan, D. (1989). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning, 44,* 493-527

Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51-78). Oxford: Oxford University Press.

Ohta, A.S. (2001). *Second language acquisition processes in the classroom: learning Japanese*. Mahwah, NJ: Lawrence Erlbaum.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes and outcomes? *Language Learning, 44*, 493-527.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on input comprehension. *TESOL Quarterly, 21*, 737–58.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System, 2,* 363-74.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research, 5*, 29–53.

Storch, N. (2002a). Patterns of interaction in ESL pair work. *Language Learning, 52*,119–58.

Storch, N. (2002b). Relationships formed in dyadic interaction and opportunity for learning. *International Journal of Educational Research, 37,* 305–22.

Storch, N. (2005). Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing, 14,* 153-73.

Storch, N. (2005). Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing, 14,* 153-73.

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research,11,* (2), 143-159

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-82). Cambridge: Cambridge 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: OUP.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: two adolescent
French immersion students working together. *Modern Language Journal, 82,* 320–38.

Swain, M., & Lapkin, S. (2002). Talking it through: two French immersion learners' response to reformulation. *International Journal of Educational Research, 37,* 285–304.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002) Peer-peer dialogue as means of second language learning. *Annual Review of Applied Linguistics, 22,* 171–85.

Van Lier, L. (2000). From input to affordance: Socio-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245-59). Oxford: Oxford University Press.

Villamil, O. S., & de Guerrero, M. C. (1996). Peer revision in the L2 classroom: social-cognitive activities, mediating strategies,

and aspects of social behavior. *Journal of Second Language Writing, 5,* 51-75.

Villamil, O. S., & de Guerrero, M. C. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics, 19,* 491-514.

Vygotsky, L. S. (1978). *Mind and society. The development of higher psychological processes.* Cambridge, Mass: Harvard University Press.

Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (ED.), *The concept of activity in Soviet Psychology* (pp. 144-188). Armonk, NY: M.E. Sharpe.

Vygotsky, L. S. (1986). Thought and language. Cambridge, Mass: MIT Press.

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners. *Language Teaching Research 11,* 121-142.

Webb, N. (1989). Peer interaction and learning in small groups. *International Journal of educational Research, 13,* 21-39.

Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press

Wells, G. (1999). Using L1 to master L2: A response to Anton and DiCamilla's "Socio-cognitive functions of L1 collaborative interaction in the L2 classroom". *The Modern Language Journal, 83,*(2), 248-254.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard
University Press.

Wertsch, J. V. (1991) *Voices of the mind: A sociocultural approach to mediated action.* Cambridge, Mass: Harvard University Press.

Appendices

*Appendix **A:** FCE Writing Paper*

Part 1

You must answer this question

1 You have decided to do the same summer job that you did last year. You want to persuade your English-speaking friend, Jan, to do the job as well.

Read the advertisement for the job and also read your notes. Then, using all your notes, write your letter persuading Jan to join you.

HELP NEEDED THIS SUMMER
Every summer we run an international
camp for 8 to 12-year-old children.
(Beautiful place …)

You will help with:

Outdoor activities
Entertainment
Sports competitions
Children's games

Free accommodation (Tental)
Free food
Only 5 hours' work a day!

Write a letter of between 120 and 180 words in an appropriate style on the opposite page. Do not write any postal addresses.

Part 2

Write an answer to one of the questions 2-5 in this part. Write your answer in 120-180 *words* in an appropriate style on the opposite page. Put the question number in the box at the top of page 5.

2 the school where you learn English has decided to buy some videos in English. You have been asked to write a report for the Principal,

suggesting what kinds of videos the school should buy. In your report, you should also explain why students at the school will enjoy these videos. Write your report.

3 you have seen the advertisement in *The Traveler* magazine and have decided to

THE TRIP OF LIFETIME

WE ARE LOOKING FOR EIGHT PEOPLE FROM DIFFERENT COUNTRIES TO SPEND THREE MONTHS ON A SAILING TRIP AROUND THE WORLD. FREE OF CHARGE. NO SAILING EXPERIENCE IS NEEDED. TRAINING WILL BE GIVEN.

WRITE AND TELL US WHU YOU ARE THE PERSON WE ARE LOOKING FOR AND EXPLAIN HOW YOU WOULD BENEFIT FROM THIS EXPERIENCE.

Write a letter of application. Do not write any postal addresses.

4 You have had a class discussion on food and eating habits. Your teacher has now asked you to write a composition giving your opinion on the following statement.

Young people are eating a less healthy diet nowadays than their grandparents did.

5 Answer one of the following two questions based on your reading of one these set books. Write the letter (a) or (b) as well as the number 5 in the question box, and the ti9tle of the book next to the box. Your answer must be about one of the books below.

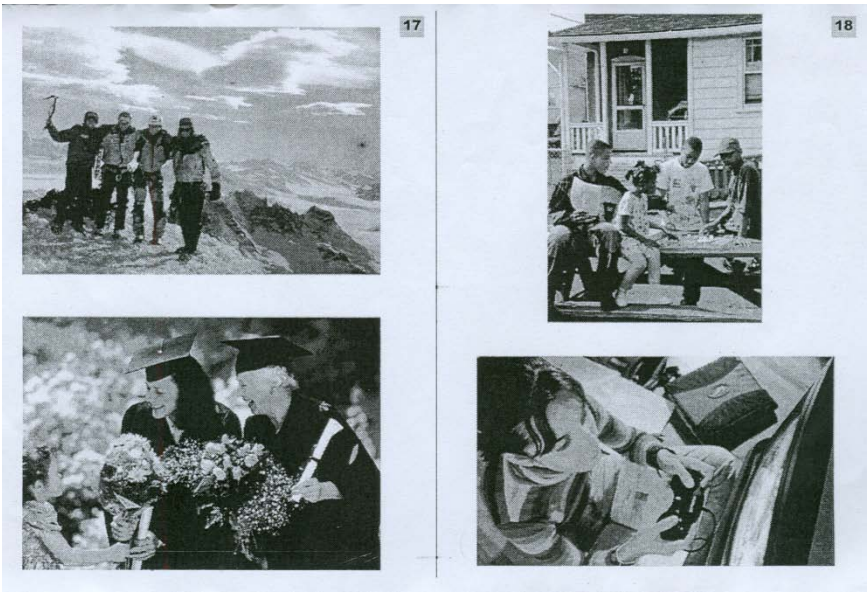
Best Detective Stories of Agata Christie – Longman Fiction
A tale of Two Cities – Charles Dickens
Animal Farm – George Orwell
Wuthering Heights – Emily Bronte
More Tales from Shakespeare – Charles and Mary Lamb

Either (a) your college magazine is looking for articles on the qualities of good stories. Write an article, briefly describing the beginning of the book or one of the short stories you have read, and explaining why this beginning made

you want to read the rest of the story.

Or (b) in your opinion, which character changes most in the book or one of the short stories `you have read? Write a composition, explaining your views.

Appendix B



Appendix C: Four Table Description Tasks

Table 1

The information about ten students from different majors who study in Tabriz Islamic Azad University and have applied for a loan from the university. Write a short composition to summarize the information

NO	sex	Age	Field of Study	City
1	F	22	Persian Literature	Marage
2	F	20	Computer	Orumie
3	M	23	Sience	Marand
4	F	19	Computer	Tehran
5	M	22	Art	Terhran
6	F	20	Science	Marand
7	F	21	English	Sarab
8	M	24	Persian Literature	Orumie
9	F	20	Art	Tehran
10	F	19	Sience	Marage

Table 2

The rate of car accidents in the four seasons in the year 1384 in Tabriz. Work in pairs and summarize the information presented in this table.

Season	Number of accidents
Spring	289
Summer	268
Autumn	310
Winter	401

Table 3

English test scores of five university classes. Work in pairs and compare the results. Write a short composition to summarize the results.

The Field of Study	The English Test Mean Score
Science	65.23
Civil Engineering	66.32
Medicine	72.28
Art	49.67
Literature	40.32

Table 4
The Pre-university students' preferences to study in different fields at the university. Write a composition to summarize the information in the table.

The Field of Study Preference	The Percentage of the Students
Medicine	87.5
Art	46.78
Science	60.56
Literature	31.24
Engineering	76.90

Appendix D: Topics for Free Compositions

1) Work in pairs to discuss the problem that the quality of education is not improving with the same pace as the number of students increase each year. Write a composition about the probable outcomes of this mismatch.

2) Work in pairs to write a composition about the increasing gap between the poor and the rich in our society.

3) Work in pairs and talk about the main problems that English students may have in our country.

4) Write a composition about the most serious threats to the present-day man on the earth.

Appendix E: TEEP Attribute Writing Scales (Weir, 1990, reproduced in Weigle, 2002)

A. Relevance and adequacy of content

0. The answer bears almost no relation to the task set. Totally inadequate answer.
1. Answer of limited relevance to the task set. Possibly major gaps in treatment of topic and/or pointless repetition.
2. For the most part answers the tasks set, though there may be some gaps or redundant information.

3. Relevant and adequate answer to the task set.

B. Compositional organization

0. No apparent organization of the content.
1. Very little organization of content. Underlying structure not sufficiently controlled.
2. Some organizational skills in evidence, but not adequately controlled.
3. Overall shape and internal pattern clear. Organizational skills adequately controlled.

C. Cohesion

0. Cohesion almost totally absent. Writing so fragmentary that comprehension of the intended communication is virtually impossible.
1. Unsatisfactory cohesion may cause difficulty in comprehension of most of the intended communication.
2. For the most part satisfactory cohesion although occasional deficiencies may mean that certain parts of the communication are not always effective.
3. Satisfactory use of cohesion resulting in effective communication.

D. Adequacy of vocabulary for purpose.

0. Vocabulary inadequate even for the most basic parts of the intended communication.
1. Frequent inadequacies in vocabulary for the task. Perhaps frequent lexical inappropriacies and/or repetition.
2. Some inadequacies in vocabulary for the task. Perhaps some lexical inappropriacies and/or circumlocution.
3. Almost no inadequacies in vocabulary for the task. Only rare inappropriacies and/or circumlocution.

E. Grammar

0. Almost all grammatical patterns inadequate.
1. Frequent grammatical inaccuracies.
2. Some grammatical inaccuracies.
3. Almost no grammatical inaccuracies.

F. Mechanical accuracy I (punctuation)

0. Ignorance of conventions of punctuation.
1. Low standard of accuracy in punctuation.
2. Some inaccuracies in punctuation.
3. Almost no inaccuracies in punctuation.

G. Mechanical accuracy II (spelling)

0. Almost all spelling inaccurate.
1. Low standard of accuracy in spelling.
2. Some inaccuracies in spelling.
3. Almost no inaccuracies in spelling