

Journal of English Language
Teaching and Learning
Tabriz University
No. 17, 2016

The Influence of Collaboration on Individual Writing Quality: The Case of Iranian vs. Malaysian College Students*

Parviz Ajideh

Professor, Tabriz University, Iran

Gerhard Leitner

Professor, Free University, Berlin, Germany

Seyed Yasin Yazdi-Amirkhiz**

Research Fellow, University of Malaya (UM), Malaysia
(Corresponding Author)

Abstract

This study purported to comparatively investigate the influence of collaborative writing on the quality of individual writing of four female Iranian and four female Malaysian students. The first semester students at a private university in Malaysia, who were comparable in terms of age, gender, study discipline, and language proficiency, were divided into two Iranian and two Malaysian dyads. The dyads performed collaborative writing tasks for 15 sessions; after three consecutive collaborative writing sessions, each participant was asked to individually attempt a writing task. Both collaborative and individual writing tasks comprised isomorphic graphic prompts (IELTS Academic Module task 1). Writing quality of the five individually-produced texts during the study was rated in terms of *task achievement* (TA), *cohesion/coherence* (C/C), *grammatical range/accuracy* (GR/A), and *lexical resources* (LR). The findings indicated a hierarchy of development in TA and C/C among all the students, while LR showed minor improvement only among three of Malaysian students, and GR/A barely exhibited any progress among everyone. Intermittent progressions and regressions were also discerned in the trajectory of their writing development. The findings are discussed in the light of the socio-cultural and emergentist perspectives, the typology of tasks used as well as the role of the participants' level of language proficiency.

Keywords: Collaborative writing; Writing quality; Individual writing.

* Received date: 2016/02/05 Accepted date: 2016/06/15

** **E-mail:** nasser_yazdi@yahoo.com

Background of the Study

Collaborative pedagogies have gained prominence in English language education globally over the past two decades. Epistemology (Gibbons, 2006), psycholinguistic arguments (Vygotsky, 1978) as well as empirical evidence (e.g., Kuiken & Vedder, 2002; Storch, 2002, 2005; Swain, 2010) support this mode of learning. Drawing upon the theoretical position of social constructivists like Bakhtin (1981), Halliday (1978), Vygotsky (1978) who all discuss language and learning as processes of meaning-making and social *activity*, the advocates of a dialogic view of writing perceive writing as a socio-contextual phenomenon and support the essential role of conversation and consultation in collaborative writing. Dialogic engagements are viewed as instrumental in the planning, production, and revision of texts (Bruffee, 1995).

A review of the literature reveals that collaborative writing pedagogies have emerged within two major frameworks, i.e., writing conferences (teacher-student) and peer-to-peer collaborative writing. Writing conferences well signify the social constructivist paradigm in writing pedagogies as per their element of Vygotsky's notion of Zone of Proximal Development (ZPD) which originally described the novice-expert collaboration, or as Storch (2002) refers to it as "asymmetrical" grouping. Despite the mention of merits for writing conferences such as that teachers have an "unparalleled opportunity to provide targeted, individualized instruction" (Weissberg, 2006, p. 261), empirical research examining the writing conference discourse has indicated that, influenced by their predominantly teacher-centered scope or transmission pedagogy nature, potentials of conferences do hardly present themselves (Freedman, 1987). The most highlighted criticisms of the conferences include learners' socio-cultural inhibitions in relation to informal engagement with teachers, contesting authority (Goldstein & Conrad, 1990), the overtly authoritative role adopted by teachers within conference settings (Wong, 1988), and the proclivity of the teachers to impose their dominant interpretive frameworks in asymmetrical interactions (Ulichny & Watson-Gegeo, 1989). However, the symmetrical power relation constitutes the salient feature of peer-to-peer collaboration. Researchers have foregrounded the significance of the symmetrical structure of power distribution in groups to the extent

that they have raised it as a strong determinant of successful collaborative development, arguing that the symmetrical configuration of groups in terms of power relations can facilitate the dynamics of interactional engagements and affective relations of the peers (Norton & Toohey, 2001). Neo-Vygotskians (e.g., Anton & DiCamilla, 1998; Kowal & Swain, 1994; Ohta, 2000; Swain & Lapkin, 1998) have likewise reported merits in symmetrical collaboration, where peers possess proximal language competences. For instance, Kowal and Swain (1994) observed that oral discourse of students on a dictogloss task increased in reciprocity.

Pertinent literature also reveals that peer-to-peer interactions have engaged in collaborative writing tasks employing two distinct frameworks: one-and-three-stage collaborative writing. One-stage collaborative writing only involves planning (Blanton, 1992), or composing (Hirvela, 1999) or revision (Williams, 2002). O'Brien (2004) and Storch (2005) agree that peer revision, which constitutes a form of collaborative writing in which "students share and comment on drafts of each other's papers" (Hirvela, 2004, p. 160), is the most common and frequently practiced peer group method of writing. Some studies have particularly looked into the effect of peer revision on text quality. A study by Hedgcock and Lefkowitz (1992) indicated a positive influence of peer revision: those EFL university students who were involved in peer-revision sessions got higher scores than their counterparts who received only written feedback from teachers. Villamil and Guerrero (1998) attributed the longer second drafts produced by 54 ESL students to the effect of peer revision activities. Berg (1999) added the variable of training to the peer revision, comparing the quality of the written texts of two groups of students: trained vs. untrained. The study found a significant positive impact of training on the quality of revisions made by the trained group.

Compared to one-stage collaborative writing and particularly peer review studies, the number of studies investigating the effect of collaboration over the entire writing process is limited. A brief review of salient studies is made here. Storch (1999) found that collaboratively-performed tasks were overall more accurate than individually-completed tasks. In Storch's (2005) study, out of 23 ESL undergraduate students at a university in Australia, 18 students chose to work

collaboratively (in pairs), and five of the students opted to work individually. She found that the jointly-produced scripts had shorter but more linguistically complex and grammatically accurate forms than the individually-written texts. In a similar study, Storch and Wigglesworth (2007) also found similar results when comparing the writing of pairs and individuals. Wigglesworth and Storch (2009) compared texts produced by 48 pairs of advanced students of English with texts produced by 48 students working individually, finding no statistically significant differences between the texts in terms of fluency and complexity. However, jointly-produced texts indicated higher grammatical accuracy than the texts produced by individuals. Shehadeh (2011) investigated the effectiveness of collaborative writing in two intact English classes in the EFL context of United Arab Emirates, finding an overall significant effect on writing, with emphasis on content, organization, and vocabulary, but on grammar or mechanics of writing (i.e., spelling, punctuation, etc.). More recently, Watanabe's (2014) study found that compared to jointly-produced texts (i.e., paired writing), individual writing further promoted the fluency of the texts, but not any greater communicative quality or accuracy.

Overall, as Storch (2011) has pointed out, despite a wide theoretical support, "the use of collaborative writing tasks in L2 classes, to date, seems relatively limited" (p. 275). Even the settings adopting collaborative modes of writing have mostly implemented compartmentalized strategies in that the collaborative practices have either occurred in the beginning stage (oral brainstorming) (Blanton, 1992), in collaborative composing (Hirvela, 1999), or with peer revision groups (Williams, 2002). For instance, the use of peer revision, which is the most prevalent mode of collaborative writing (O'Brien, 2004; Storch, 2005), has been criticized on the grounds that during peer revisions students tend to focus on lexico-syntactic level issues, rather than on important revising issues (Lockhart & Ng, 1995; Nelson & Carson, 1998; Villamil & Guerrero, 1996). Besides this, the focus of the learners in peer review activities is on the finished product, not on the process of writing (Storch, 2005). Numerous scholars (Daiute, 1986; Dobao, 2012; Ede & Lunsford, 1990; Storch, 2011, 2013; Wells, Chang, & Maher, 1990) support and argue for the collaborative engagement of students throughout the entire writing process (i.e.,

throughout brainstorming, composing and revising stages). It has even been argued that collaborative writing should be distinguished from other forms of tasks such as peer editing and peer feedback in that they tend to compartmentalize the writing process and limit the writers' interaction to only one stage of writing and thus, they do not share co-ownership of the joint text (Storch, 2013). Most of the previous empirical studies (e.g., Kim, 2008; Storch, 2002; Swain, 2000, 2010; Swain & Lapkin, 1998) have, in fact, utilized dictogloss and text reconstruction tasks, which are far from being the "authentic" composition tasks (Donato, 1988).

Moreover, earlier studies have mostly collected and analyzed data from a single writing session, evaluating and comparing the accuracy, complexity, and rhetoric organization of jointly-written texts with individually-composed pieces (e.g., Dobao, 2012; Kuiken & Vedder, 2002; Storch, 1999, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). As Shehadeh (2011) rightly puts it, in most of the existing studies, the pedagogically important issue of eventual effect of collaborative writing on the quality of individual L2 writing has been taken for granted. Also, most of the prior studies (e.g., Kuiken & Vedder, 2002; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009) have focused on second language (SL) contexts, and scant research has been done on FL contexts (e.g., Shehadeh, 2011). The point to be made here is that notwithstanding a blurred line of demarcation between ESL and EFL in the current mainstream L2 education, there yet exist certain features which render delineation of EFL and ESL realms unavoidable. As for the present study, the socio-contextual pervasiveness of the English language in the two contexts of Iran and Malaysia, where the instructional socialization of the participants of the present study took place, are conspicuously distinct. English in Malaysia has a very strong presence in the context of society and as Baker (2008) puts it "is used as an institutionalized additional language" (p. 132). But in Iran English hardly has a near socio-contextual function that it does in Malaysia, and is hence considered a foreign language (Yarmohammadi, 2005). English learning histories of the participants were likewise tangibly distinct (see Appendix). Thus, drawing upon Grabe and Kaplan's (1996) theorization that different cultural backgrounds and pedagogical

socialization could affect the collaborative nature and outcomes of interactions, the current study also aims to explore the transpiring of concurrent implementation of collaborative writing among Iranian and Malaysian students. Based on the foregoing, we pose the following research question so to guide the study:

1. How does collaborative writing influence the individual writing quality of female Malaysian and Iranian learners?

Methodology

Participants and setting

The criterion-based sampling technique was employed and eight first-semester undergraduate female students (i.e., four Malaysian-Chinese and four Iranian students) were selected at a private university in Kuala Lumpur, Malaysia. Factors that can influence interaction and group dynamics include the group members' language competence (Grabe & Kaplan, 1996; Villamil & Guerrero, 2006), sex (gender enactment) (Chavez, 2000; Gass & Varonis, 1986), and cultural background and pedagogical socialization (Grabe & Kaplan, 1996). Thus, the criteria that each of the participants was expected to fulfill was: (1) they had proximal language proficiency; (2) they were all female; and (3) they were all from a common discipline. In order to further measure the participants' writing ability, in addition to their most recent writing proficiency test result (IELTS), another writing proficiency test adapted from IELTS was also administered to them. All participants reproduced their scores of 6 (out of 9) in the IELTS Academic Module writing. The eight participants were divided into two Iranian dyads (A & B) and two Malaysian dyads (C & D). The pseudonyms used for the participants were: Azadeh and Sadaf (Dyad A), Negar and Niloufar (Dyad B), Mei and Tang (Dyad C), and Gin and Wai (Dyad D).

As for the rationale behind adopting a dyadic configuration rather than groups in the present study, research has shown that in groups one or two members dominate the group, which can impede the participation of others (Kagan, 1994). Some scholars have suggested dyads as the most appropriate grouping for language classes (Doughty & Pica, 1986; Kowal & Swain, 1994). Peacock (1998) found that "learners worked significantly harder (that is, spent significantly more time on-task) in pairs than in groups of three" (p. 37). Shehadeh (2004) substantiated that dyadic interaction provides "quantitatively more

opportunities for Modified Output than group interaction” (p. 351). Wang (2009) has supported the symmetrical configuration of dyadic structuring on the grounds that it may enhance individual accountability.

Pilot study

A pilot study was conducted to ensure the quality of the study, to establish the feasibility of the research, as well as to expose any probable challenges earlier than the main study. The pilot study also sought to trial the IELTS Academic Module (AM) Task 1 when it was done collaboratively. Two Malaysian undergraduate students (i.e., a Malaysian dyad) from a public university and two undergraduate Iranian students (i.e., an Iranian dyad) from a private university in Kuala Lumpur functioned as the participants of the pilot study. Students in each dyad were asked to collaboratively perform a sample writing task which had been randomly-selected from IELTS AM Task 1. Adhering to IELTS testing system requirements for Task 1, each dyad had 20 minutes to complete the task. Upon completion of the first task by each dyad, the researcher attempted to elicit opinions concerning the study design from the participants. All participants expressed their content with the task-completion procedure. However, they complained about the allocated task completion time, requesting an extension. Considering students’ complaints in reference to time limit as well as suggestions by Storch (1999, 2005) that collaborative writing requires more time than individual writing, students received an additional ten minutes to complete the tasks. They did two more collaborative writing tasks. Despite contributing to rethinking collaborative task completion time, the pilot study did not contribute in significant ways to the main study.

Data collection procedure

The study collected and collated data during semester vacations, as it aimed to observe the students at times during which they had minimal exposure to academic English, and consequently there was little concern with the learning process the participants might have gone through as a significant factor. Data collection lasted approximately eight weeks. Each dyad selected a time convenient for interaction; therefore, data collection occurred at times differing for each dyad. Data

collection procedures for each dyad comprised of five cycles, where each cycle consisted of three consecutive collaborative writing sessions, followed by performing one individual writing task. In each collaborative writing session the researchers presented each of the four dyads with identical graphic prompts (IELTS AM task 1) requesting that the participants collaboratively perform the tasks in a maximum of thirty minutes, and in a minimum of one-hundred and fifty words. In each cycle, at completion of the third session of collaborative writing, all participants received identical writing tasks which they had to perform individually, and within a 20-minute time limit. It is a truism that each participant produced five individually-completed writing tasks in the course of the study.

Validity and reliability

The wide scale usage of English language tests such as the IELTS globally, together with the significance of the results from these tests toward critical decisions on test-takers' lives (e.g., tertiary-level admission) has obviously influenced the consideration toward technical and professional standards in these tests. IELTS is currently regarded as a test with an acceptable level of validity and reliability. Uysal (2010) enumerates parameters contributing to the validity and reliability of the IELTS writing test. With reference to validity, the tasks usually represent what students are likely to be assigned to do in authentic contexts and "the topics and the context of language use, which might introduce a bias against any group of candidates of a particular background, are avoided" (p. 316) in the test. The analytic scale contributes to higher reliability in the marking scheme. Uysal (2010) and Weigle (2002) state that the reliability increases through professional training of raters as well as double-rating practice. However, it should be noted that in IELTS centers double-rating does not occur for every script, but rather for the ones with uneven overall profiles.

The present study rated each participant's five individually-produced scripts using IELTS rating-scale descriptor (Public version). The raters included one professional IELTS examiner, and an experienced rater (a PhD student of TESL with over ten years of experience in teaching IELTS). The high expenses of rating by IELTS professional examiners was the main reason behind choosing the

second rater from among PhD students of TESL. The second rater was also compensated for his service. It was also envisioned that a third rater be used to adjudicate in cases of marked discrepancy. Tedick's (1990) suggestion grounded the scoring procedure that "[m]ultiple independent systematic scoring involves having two different readers score each essay independently. A one-point difference between scores is acceptable; a difference of two or more points necessitates a third reading, which resolves the discrepancy" (p.140). Drawing upon Tedick's suggestion, when the difference between the scores was one point, the scores were averaged. In fact, at no time did rater scores differ by more than one-point, eliminating the need for a third rater.

Findings and Discussion

In this section, the analysis and discussion of the salient findings are presented. Each participant's individual writing quality across the study is shown by the tables 1-8 below.

Table 1. Azadeh's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5	5.5	6.5	6.5	6
TASK 2	6	6.5	6	6	6
TASK 3	6	6	6.5	6	6
TASK 4	6	6	6	6	6
TASK 5	6.5	6	6	6	6

Table 2. Sadaf's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5	5.5	6	6	5.5
TASK 2	5.5	6	5.5	5.5	5.5
TASK 3	5.5	5.5	5.5	5.5	5.5
TASK 4	6	5.5	5.5	6	6
TASK 5	6	6	5.5	6	6

Table 3. Negar's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	6	6	6	6	6
TASK 2	6	6.5	6	6.5	6
TASK 3	6	6	5.5	6	6
TASK 4	6.5	6.5	6	6	6
TASK 5	6.5	6	6	6	6

Table 4. Niloofar's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5.5	6	5.5	6	6
TASK 2	5.5	5.5	5.5	6	5.5
TASK 3	6	6	6	5.5	6
TASK 4	6	5.5	5.5	6	6
TASK 5	6	6	5.5	6	6

Table 5. Mei's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5.5	5.5	6	5	5.5
TASK 2	5.5	5.5	6	6	6
TASK 3	6	6	6	5.5	6
TASK 4	6	6	6	5.5	6
TASK 5	7	6	6.5	5	6

Table 6. Teng's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5.5	5	6	6	5.5
TASK 2	5	5.5	6	5.5	5.5
TASK 3	6	5	6	6	6
TASK 4	6	5	6	5.5	5.5
TASK 5	6	6	6.5	5	6

Table 7. Gin's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	5.5	5.5	5.5	5.5	5.5
TASK 2	5.5	5	5	5.5	5.5
TASK 3	6	5.5	5.5	5	5.5
TASK 4	6	5.5	5.5	5.5	5.5
TASK 5	6	5	6	5.5	5.5

Table 8. Wai's writing quality across five individual tasks

	TA	C/C	LR	GR/A	TOTAL
TASK 1	6	5.5	5.5	6	6
TASK 2	5.5	5.5	5.5	6	5.5
TASK 3	6.5	5	5.5	5.5	6
TASK 4	6.5	6	5	5.5	6
TASK 5	6.5	6	5.5	5.5	6

The results of the five individually-performed tasks for Malaysian and Iranian participants (see tables 1-8) indicate that participants' involvement in collaborative writing sessions influenced quality of individual writing. Judging by the IELTS Handbook's (2002) progress

benchmark that “individuals can take up to 200 hours to improve by one IELTS band” (p. 22), the overall gains obtained by the participants sound almost substantial. Although the total band scores appear unchanged for six of the participants, the individual band scores indicate an improvement. Since the total band scores represent the average value (the sum of the individual scores divided by four), they (i.e., the total band scores) sometimes fail to reflect small changes in individual band scores. For example, the total band score of a candidate with individual scores of 6 in each of the four individual band scores (TA), (C/C), (LR), (GR/A) would be equal to the total band score of a candidate whose individual task scores are 6, 6, 5.5, 5.5. Therefore, due to the fact that individual band scores represent a higher degree of accuracy, they have been addressed and adopted in the present study.

Overall, TA followed by C/C represented the most visible similarity between Iranian and Malaysian participants’ writing performance. Moreover, Iranian and Malaysian participants indicated a similar lack of improvement (at times regression) on the dimension of GR/A. With reference to LR, unlike Iranians whose lexical range score did not show any improvement, three Malaysian participants showed minor improvement in their scores on their final individual task. For instance, whereas Azadeh and Mei raised their scores by one and half points in TA component, Gin and Niloofar only managed to raise their scores by 0.5 point. C/C constituted the second most positively affected area for many participants. However, unlike TA, where the extent of improvement showed considerable change, C/C showed small change, and also included regression. We may explain the patterns in TA and C/C through suggesting the presence of situated practice. Task achievement (TA) in IELTS TASK 1 associates with the test-takers’ competences in selecting the most important information in the pictorial data, where C/C may associate with their ability to connect ideas. With respect to TA and C/C, given the use of isomorphic tasks, it could be hypothesized that repeated attempts with similar tasks may have enhanced student dexterity in selecting appropriate information (the most important features) from the isomorphic tasks and organizing them in an acceptable manner. Apart from the likely ‘practice’ effect, Crosthwaite’s (2011) study has also suggested that collaborative

engagements might positively impact on cohesion and coherence of non-native speakers' narrative discourse.

Unlike TA and C/C, the GR/A not only remained unaffected for almost all participants, but back and forth patterns emerged across the tasks. LR did not indicate substantial improvement either; the lexical range score of the Iranian students did not show any improvement and three of Malaysian participants indicated a minor improvement in their final individual writing task. With respect to GR/A and LR, the lack of noticeable improvement indicates inconsistency from the expectation that the collaborative engagements, in view of socio-cultural perspectives of language development, would affect increased competences in grammatical accuracy and lexical maneuvering. It has been found that collaborative writing could significantly affect the grammatical accuracy of writing (Kuiken & Vedder, 2002; Storch, 2005; Storch & Wigglesworth, 2007; Watanabe, 2014; Wigglesworth & Storch, 2009), although Shehahdeh's (2011) study suggested otherwise. The results of the present study also indicate inconsistency with other studies (Kim, 2008; Shehadeh, 2011) in that their findings favor collaborative writing for the enhancement of vocabulary knowledge. These inconsistencies may have to do with the respective pedagogical enculturations of the participants. According to Manglesdorf (1992), owing to the teacher-centeredness of Asian cultures, Asian learners usually look down on the feedback and comments provided by their peers. In Malaysia, Maarof, Yamat, and Lili (2011) found out that Malaysian learners "...placed their teachers in a higher position in giving feedback to their essays compared to their peers" (p. 32). In Iran, Rabiee (2008), who conducted a peer response study among 60 Iranian students, found that using dyadic peer response groups in writing classes indicated no positive effect on students' revising and writing outcomes. Her study also reported problems with and resistance by the participants towards peer revision. She attributed the findings to the socio-cultural specificities of Iranian culture. According to her, "...in most teacher-centered contexts, like Iran, since the students are not familiar with a collaborative, learner-centered environment, such an activity may be resisted by them and hence negative views may be developed" (p. 14).

Other explanation for the difference in the results of the present study with previous studies may be the use of different measures of grammatical accuracy in this study. Drawing on Shehadeh's (2011) explanation for the lack of significant impact of collaborative writing on the grammatical accuracy of his students, we may likewise argue that the same explanation may sound plausible for the present study as well. He partly attributed the difference in the results of his study with those of the studies by Storch (2005) and Storch and Wigglesworth (2007) to the use of different measures of grammatical accuracy, arguing that:

Storch (2005) and Storch and Wigglesworth (2007) used error-free clauses as measures of grammatical accuracy, that is, calculating the proportion of error-free clauses in relation to the total number of all clauses used in the text. The current study [Shehadeh's study], on the other hand, used global scales, that is, criteria based on the rater's judgment of the student's performance on a particular component or area of the text like content, organization, grammar etc. (p.11)

Furthermore, the small scale nature of the present study may have probably influenced the results obtained. Conducting a study with a larger sample size could have presented a different picture of the influence of collaborative learning.

Results of the present study might also become justified based on one main explanation and one minor explanation. The former might be related to the participant proficiency, whereas the latter may have something to do with the type of tasks employed in the study. The participants' somehow higher-intermediate to advanced proficiency level in L2 could be considered as one of the important factors affecting the quality of interaction between the learners. Based on the Common European Framework of Reference Levels (established by the Council of Europe), IELTS band scores 5, 5.5 and 6 correspond to the upper-intermediate level. SLA research indicates that feedback can more effectively assist students of intermediate proficiency than with beginning or more advanced students (Carroll, 1995, 2001; Lyster & Ranta, 1997). Feedback can less effectively assist beginning learners in that their low proficiency limits their linguistic repertoire from which

to appropriate feedback. Similarly, feedback can also prove less effective with advanced learners, partly influenced by that these learners might concern themselves less with linguistic accuracy if they manage to convey the semantics of purpose, and partly influenced by that their language affordances may have prematurely leveled off at a certain level (Carroll, 1995; Long, 1996). Thus, such a scenario may have eventuated with the participants' proficiency in the present study. In addition to the parameter of language competence and proficiency, the type and content of the tasks in the study may have also influenced the results obtained. As said earlier, the tasks employed were isomorphic, and the actual tasks which the learners performed were naturally not conceptually difficult, challenging or new to them, therefore there was no need for them to work on language issues in order to achieve task requirements successfully.

The noticeable number of fluctuations in the students' English writing performance across time was presumably the most striking aspect of their writing. The non-linear, unpredictable and chaotic nature of language development in dynamic complex systems (e.g., Larsen-Freeman, 1997; Cameron & Larsen-Freeman, 2007) might probably be the most plausible explanation for the erratic writing performance of the participants. From the perspective of dynamic complex systems or an "ecological" (Van Lier, 2000, 2004) view of language education, language development does not emerge in a linear process, but rather, as a semiotic emergence (Lantolf, 2006). According to Larsen-Freeman (2006), an emergentist perspective provides a useful way of conceptualizing L2 development systems. Based on this perspective, "...[language] development is not discrete and stage-like but more like the waxing and waning of patterns...certain aspects of behavior are progressive, others, regressive" (p. 590). The findings of the present study converge with those of Larsen-Freeman's (2006) longitudinal study of five Chinese participants' oral and written production in terms of complexity, accuracy and fluency where learners exhibited significant variation and fluctuation in their performances over time. Larsen-Freeman (2009) described the five students' performances in her subsequent paper that "[s]ometimes they [graph lines illustrating language development] went up; sometimes they went down. Some

learners finished their six-month course worse off with regard to a particular CAF dimension than when they had started!” (p. 586).

Implications of the Study

The findings of this study may have some theoretical and pedagogical implications. Given that socio-cultural perspective of language learning underlies this study, one theoretical implication of the study is that the findings reconfirm the efficacy of collaborative writing in improving the learners’ quality of individual writing. It is also adding to the previous attempts to extend the socio-cultural perspective of language learning from the traditionally predominant spoken discourse to written discourse (collaborative writing) (e.g., Kuiken & Vedder, 2002; Shehadeh, 2011; Storch, 2005; Storch & Wigglesworth, 2007). Furthermore, the positive results obtained could be partly attributed to the framework and configuration of collaborative writing that was employed in the present study; the writing process was not compartmentalized, the focus was on the entire writing process, and the dyadic structuring was symmetrical in terms of power relations. It could be argued that grouping the participants of the dyads over the entire writing process provided them with an opportunity to co-construct the texts from the beginning of the process to the end. Such a possibility may have given them a sense of ownership of what they were co-constructing in that they had latitude to create a meaning that was their own, within a micro-social context defined by the peers themselves, while being cognizant of their shared responsibility of working for a common goal (co-producing a text). Another theoretical contribution of this study is that collaborative writing may positively affect the *task achievement* and *cohesion/coherence* aspects of writing ability. This finding extends the boundaries of the benefits of collaborative writing; the benefits go beyond lexico-grammatical level as addressed by previous studies (e.g., Kuiken & Vedder, 2002; Storch, 2005) and encompass *task achievement* and *cohesion/coherence* facets of writing proficiency as well. Last but not least, the unpredictable patterns of regressions and progressions in the participants’ writing performance provide another tangible piece of evidence in support of the emergentist conceptualization of language learning process, which can in turn lead us (i.e., teachers) to develop more realistic and down-to earth expectations about writing development ‘trajectory’ of our students.

Acknowledgements

We would like to thank Professor Azirah Hashim, dean of the Humanities Research Cluster at University of Malaya, for all her generous help and support as well as Dr. Dimitrios Michael Hadzantonis for reading the earlier draft of this manuscript and providing his insightful comments on different aspects of the study.

References

- Anton, M. & DiCamilla, F. J. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 54(3), 314-342.
- Baker, W. (2008). A critical examination of ELT in Thailand. *RELC Journal*, 39(1), 131-146.
- Bakhtin, M. (1981). *The dialogic imagination: Four essays*. Austin: University of Texas Press.
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241.
- Blanton, L. L. (1992). Talking students into writing: Using oral fluency to develop literacy. *TESOL Journal*, 1(4), 23-26.
- Bruffee, K. A. (1995). Peer tutoring and the "conversation of mankind." *Landmark Essays: Writing Centers*, 87-98.
- Cameron, L. & Larsen-Freeman, D. (2007). Complex systems and applied linguistics. *International Journal of Applied Linguistics*, 17(2), 226-240.
- Carroll, S. E. (1995). The irrelevance of verbal feedback to language learning. In L. Eubank, L. Selinker & M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford* (pp. 73-88). Amsterdam: John Benjamins.
- Carroll, S. E. (2001). *Input and evidence: The raw material of second language acquisition* (Vol. 25). Amsterdam: John Benjamins Publishing Company.
- Chavez, M. (2000). Teacher and student gender and peer group gender composition in German foreign language classroom discourse: An exploratory study. *Journal of Pragmatics*, 32(7), 1019-1058.

- Crosthwaite, P. (2011). The Effect of Collaboration on the Cohesion and Coherence of L2 Narrative Discourse between English NS and Korean L2 English Users. *Asian EFL Journal*, 13(4), 135-166.
- Daiute, C. (1986). Do 1 and 1 Make 2? Patterns of influence by collaborative authors. *Written Communication*, 3(3), 382-408.
- Dobao, A. F. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21, 40-58.
- Donato, R. (1988). Beyond group: A psycholinguistic rationale for collective activity in second-language learning. Unpublished doctoral dissertation. University of Delaware.
- Doughty, C. & Pica, T. (1986). "Information Gap" tasks: Do they facilitate second language acquisition? *TESOL Quarterly*, 20(2), 305-325.
- Ede, L. & Lunsford, A. (1990). *Singular texts/plural authors*. Carbondale: Southern Illinois University Press.
- Freedman, S. (1987). Response to student writing: Teaching and learning. *NCTE Research Monograph Series*. Urbana, IL: National Council of Teachers of English.
- Gass, S. M. & Varonis, E. (1986). *Sex differences in NNS/NNS interactions*. Rowley, MA: Newbury House.
- Gibbons, P. (2006). *Bridging discourses in the ESL classroom: Students, teachers and researchers*. Continuum Intl Pub Group.
- Goldstein, L. M. & Conrad, S. M. (1990). Student input and negotiation of meaning in ESL writing conferences. *TESOL Quarterly*, 24(3), 443-460.
- Grabe, W. & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. New York: Longman.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Arnold.
- Hedgcock, J. & Lefkowitz, N. (1992). "Collaborative oral/aural revision in foreign language writing instruction." *Journal of Second Language Writing*, 1(3), 255-276.

- Hirvela, A. (1999). Collaborative writing instruction and communities of readers and writers. *TESOL Journal*, 8(2), 7-12.
- Hirvela, A. (2004). *Connecting reading and writing in second language writing instruction*. University of Michigan Press.
- International English Language Testing System (2002). *The IELTS Handbook, University of Cambridge Local Examinations Syndicate*. The British Council, IELTS Australia.
- International English Language Testing System (2007). *The IELTS Handbook, University of Cambridge Local Examinations Syndicate, The British Council, IELTS Australia*.
- Kagan, S. (1994). *Cooperative learning*. San Clemente, CA: Kagan Publications.
- Kim, Y. J. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92(1), 114-130.
- Kowal, M. & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness 1. *Language Awareness*, 3(2), 73-93.
- Kuiken, F. & Vedder, I. (2002). Collaborative writing in L2: The effect of group interaction on text quality. In Rijlaarsdam, Gert (Series Ed.), Ransdell, Sarah & Marie-Laure Barbier (Vol. Eds.), *Studies in writing*, 169-188. Dordrecht: Kluwer.
- Lantolf, J. P. (2006). Language emergence: Implications for applied linguistics-A sociocultural perspective. *Applied Linguistics*, 27(4), 717-728.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141-165.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579-589.

- Lockhart, C. & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45(4), 605-651.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. J. Bhatia (Eds.), *Handbook of second language acquisition* (Vol. 26, pp. 413-468). Orlando, FL: Academic Press.
- Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37-66.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom. *ELT Journal*, 46(3), 274-284.
- Maarof, N., Yamat, H., & Lili, K. (2011). Role of teacher, peer and teacher-peer feedback in enhancing ESL students' writing. *World Applied Sciences Journal*, 15 (Innovation and Pedagogy for Life Long Learning), 29-35.
- Nelson, G. L. & Carson, J. G. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing*, 7(2), 113-131.
- Norton, B. & Toohey, K. (2001). Changing perspectives on good language learners. *TESOL Quarterly*, 35(2), 307-322.
- O'Brien, T. (2004). Writing in a foreign language: Teaching and learning. *Language Teaching*, 37(01), 1-28.
- 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.
- Peacock, M. (1998). A useful and popular lesson? Comparing students working in pairs and threes in the ESL Classroom. *RELC Journal*, 29(2), 27-49.
- Rabiee, M. (2008, July). *Grouping students in peer response: The Iranian EFL students' attitudes toward dyadic grouping in writing classes*. Paper presented at the Asia TEFL International Conference, Bali, Indonesia.

- Shehadeh, A. (2004). Modified output during task-based pair interaction and group interaction. *Journal of Applied Linguistics*, 1(3), 351-382.
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), 286-305.
- Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, 27(3), 363-374.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119-158.
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153-173.
- Storch, N. (2011). Collaborative Writing in L2 Contexts: Processes, Outcomes, and Future Directions. *Annual Review of Applied Linguistics*, 31(1), 275-288.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters.
- Storch, N. & Wigglesworth, G. (2007). Writing tasks: Comparing individual and collaborative writing. In M. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157-177). London, UK: Multilingual Matters.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. *Sociocultural theory and second language learning*, 97-114.
- Swain, M. (2010). Talking-it through: Linguaging as a source of learning. In R. Batstone (Ed.), *Sociocognitive perspectives on language use/learning* (pp. 112-130). Oxford: Oxford University Press.
- Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82(3), 320-337.
- Tedick, D. J. (1990). ESL writing assessment: Subject-matter knowledge and its impact on performance. *English for Specific Purposes*, 9(2), 123-143.

- Ulichny, P. & Watson-Gegeo, K. A. (1989). Interactions and authority: The dominant interpretive framework in writing conferences. *Discourse Processes*, 12(3), 309-328.
- Uysal, H. H. (2010). A critical review of the IELTS writing test. *ELT Journal*, 64(3), 314-320.
- Van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *sociocultural theory and second language learning* (pp. 245–259). Oxford: Oxford University Press.
- Van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective* (Vol. 3): Kluwer Academic Pub.
- Villamil, O. S. & Guerrero, M. C. M. D. (1996). Peer revision in the L2 classroom: Social-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing*, 5(1), 51-75.
- Villamil, O. S. & Guerrero, M. C. M. D. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, 19(4), 491-514.
- Villamil, O. S. & Guerrero, M. C. M. D. (2006). Sociocultural theory: A framework for understanding the social-cognitive dimensions of peer feedback. *Feedback in second language writing: Contexts and issues*, 23-42.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, Q. (2009). Design and evaluation of a collaborative learning environment. *Computers & Education*, 53(4), 1138-1146.
- Watanabe, Y. (2014). *Collaborative and independent writing: Japanese university English learners' processes, texts and opinions*. Unpublished doctoral dissertation. University of Toronto, Canada.
- Weigle, S. C. (2002). *Assessing Writing*: Ernst Klett Sprachen.
- Weissberg, R. (2006). Scaffolded feedback: Tutorial conversations with advanced L2 writers. *Feedback in second language writing: Contexts and issues*, 246-265.

- Wells, G., Chang, G. L. M., & Maher, A. (1990). Creating classroom communities of literate thinkers. In S. Sharan (Ed.), *Cooperative learning: Theory and research*. New York: Praeger.
- Wigglesworth, G. & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466.
- Williams, J. (2002). Undergraduate second language writers in the writing center. *Journal of Basic Writing*, 21(2), 73-91.
- Wong, I. (1988). Teacher-student talk in technical writing conference. *Written Communication*, 5, 444-460.
- Yarmohammadi, L. (2005). ESP in iran from language planning perspective. Proceedings of the first national ESP/EAP conference (Vol. II, pp. 2-20). Tehran: The center for research and development in humanities of SAMT.

Appendix

A synoptic account of the participants' self-reported English learning histories

Malaysian and Iranian students of the present study had different English learning histories. The four Malaysian participants had a fairly similar language learning experience. They had learnt English for eleven years under Malaysia's educational system and except Wai, the other three had not attended any private language institutes. However, the Iranian students had learnt English for 7 years under Iranian educational system and had all attended private language institutes as well. As for the areas of focus in their language classes, Malaysian participants said in their English classes language teaching mainly focused on the four skills of language (reading, writing, listening, and speaking) as well as on grammar and vocabulary. However, Iranian participants unanimously stated that the focus of instruction in their English classes at schools was on grammar exercises, vocabulary and translation from English into Persian. The focus of the textbooks was also on grammar, vocabulary and reading skill. Iranian National University Entrance Exam (INUEE) was also mentioned as a factor which intensified teachers and students' attention towards grammar. It was basically argued that since the focus of the high-stakes test was exclusively on vocabulary, grammar and reading, most of the English teachers were teaching to the test and the students were likewise excessively obsessed with grammar and vocabulary.