Collaborative writing: Product, process, and students’ reflections

Neomy Storch

Department of Linguistics and Applied Linguistics, School of Languages, 5th Floor, Arts Centre, The University of Melbourne, Parkville, Victoria 3010, Australia

Abstract

Although pair and group work are commonly used in language classrooms, very few studies have investigated the nature of such collaboration when students produce a jointly written text. This study set out to investigate collaborative writing. The study was classroom based, and the participants (23) were adult ESL students completing degree courses. Students were given a choice to write in pairs or individually. Although most chose to work in pairs, some chose to work individually. All pair work was audiotaped and all completed texts collected. All pairs were also interviewed after class. The study compared texts produced by pairs with those produced by individual learners and investigated the nature of the writing processes evident in the pair talk. The study also elicited the learners’ reflections on the experience of collaborative writing. The study found that pairs produced shorter but better texts in terms of task fulfilment, grammatical accuracy, and complexity. Collaboration afforded students the opportunity to pool ideas and provide each other with feedback. Most students were positive about the experience, although some did express some reservations about collaborative writing.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Collaborative writing; Second language writing; Learners’ attitude to pair work

The use of small group and pair work in classrooms, particularly in second language (L2) classrooms, rests on strong theoretical and pedagogical bases. From a theoretical perspective, the use of small groups/pairs accords with a social constructivist view of learning. The roots of social constructivism are based on the work of Vygotsky (1978). According to Vygotsky, human development is inherently a socially situated activity.
A child’s (novice) cognitive development arises in social interaction with a more able member of society. The more able member (expert), by providing the novice with the appropriate level of assistance, stretches the novice beyond their current level towards their potential level of development. Such assistance is now commonly referred to in the literature as scaffolding. However, as a number of researchers have shown (e.g., Donato, 1994; Storch, 2002), scaffolding can also occur among peers when working in group/pair work. Thus, from a social constructivist perspective, learners should be encouraged to participate in activities which foster interaction and co-construction of knowledge. From a pedagogical perspective, the use of small group and pair work is further supported by the communicative approach to L2 instruction and its emphasis on providing learners with opportunities to use the L2.

However, the use of small group/pair work in writing classes seems quite limited. It tends to be limited to the beginning stages (brainstorming), or more commonly, to the final stages of writing—the peer review stage. In this final stage, students review each other’s written text and make suggestions on how it could be improved. A number of researchers (e.g., Ferris, 2003) have noted the benefits of such peer reviews. Foremost among these benefits is that peer reviews are a way of raising students’ awareness of audience considerations (Leki, 1993), and at the same time, they may help learners develop analytical and critical reading and writing skills (Nystrand & Brandt, 1989).

One of the drawbacks of peer reviews, however, is that the focus is often on the product of writing rather than the process of writing. In L2 contexts in particular, a number of studies (e.g., Lockhart & Ng, 1995; Nelson & Carson, 1998; Villamil & de Guerrero, 1996) have shown that when students are asked to peer review, they tend to focus on errors at the sentence and word level. Thus, the process of writing remains a private act, where writers are left to their own devices when making important decisions about their text (Hirvela, 1999).

A number of scholars, writing about developing L1 writing skills (e.g., Daiute, 1986; Wells, Chang, & Maher, 1990), argue that students should collaborate throughout the writing process. Such collaboration means that learners have joint responsibility over the production of the text. This may promote a sense of co-ownership and hence encourage students to contribute to the decision making on all aspects of writing: content, structure, and language. Ede and Lunsford (1990) refer to this as a singular text/plural authors approach.

Research findings on collaborative writing have been positive. Research conducted in L1 settings (e.g., Higgins, Flower, & Petraglia, 1992; Keys, 1994) has shown that collaborative writing is a way to foster reflective thinking, especially if the learners are engaged in the act of explaining and defending their ideas to their peers. Research conducted with L2 learners (e.g., Donato, 1988; DiCamilla & Anton, 1997; Storch, 2002; Swain & Lapkin, 1998) has shown that in the process of co-authoring, learners consider not only grammatical accuracy and lexis but also discourse. Furthermore, and depending on the kind of group/pair dynamics formed (see Donato, 1988; Storch, 2002, 2003), collaborative writing may encourage a pooling of knowledge about language, a process Donato termed collective scaffolding (Donato, 1988, 1994).

However, despite the support in the research literature for collaborative writing, as a language teacher, I have often noticed that when I ask students to work in pairs (or small
groups) on tasks which require written output, some students seem reluctant to do so. They seem to prefer to complete such tasks individually. Such observations have been reported by other L2 teachers as well (e.g., Peretz, 2003). Thus, when faced with some opposition by learners to group/pair joint writing activities, the question facing language teachers is how strongly should they encourage learners to work jointly on writing tasks. Are compositions produced collaboratively better than those produced by students writing individually? What is the nature of the writing process that pairs engage in? These are some of the questions that the present classroom-based study set out to investigate.

It should be noted that studies on collaborative writing have, by and large, focused on the attention to language and the socio-cognitive processes evident in such interactions. The nature of the writing process and of the written text produced have received scant attention. There are very few studies that have compared compositions produced collaboratively with compositions produced individually. A study by Storch (1999), which compared individual and pair performance on a range of grammar-focused exercises, including a short composition, found that exercises completed in pairs were generally more accurate than when completed individually. However, in that study the same students performed the exercises in pairs and individually.

The other issue that needs further investigation is learners’ views concerning collaborative writing. Results of studies on students’ attitudes to group/pair work in general are mixed. Some studies report that learners have predominantly positive attitudes to pair and group work (e.g., Mishra & Oliver, 1998; Roskams, 1999), while others report that learners have reservations about pair and group work (e.g., Hyde, 1993; Kinsella, 1996). However, it should be noted that most of these studies rely on surveys rather than on interviews conducted with students immediately after experiencing a collaborative activity. Moreover, these studies have examined learners’ attitudes to group/pair work in general, rather than to the activity of collaborative writing.

Thus, the current study set out to investigate the product, process, and student reflections on collaborative writing. Specifically, the study set out to compare the texts produced by learners individually with those produced by learners working in pairs. For learners working in pairs, the study also investigated the nature of the composing process. Finally, the study attempted to elicit the learners’ reflections on the activity of collaborative writing.

1. The study

The study was conducted in an ESL writing class offered for credit at a large Australian university. Two levels are available: ESL 1 and a higher level ESL 2. The study was conducted in ESL 1. Student placement in these classes is determined by a university-developed diagnostic language test. Students who are advised to enroll in these classes are considered in this context to be of intermediate language proficiency. That is, although these students have the required ESL proficiency to be accepted for study at the university (an average of 6.5 on IELTS or 233 on the computer-based TOEFL test), their scores on the diagnostic test indicate that they require additional attention to their academic writing skills and grammatical accuracy.
The task used in the study was a data commentary text. Students were given a graphic prompt and asked to compose a short (one to two paragraphs) text. The prompt showed the language proficiency of two groups of migrants (Vietnamese and Laotians) before and after coming to Australia (see Appendix A). The task was presented as class preparation for the upcoming assignment. The task was not graded, but it was collected at the end of the class and subsequently returned with feedback comments.\textsuperscript{1}

The study was conducted in two parallel intact classes. The researcher taught both classes. Data collection occurred in week 4 of the semester, thus the students were familiar with each other. Given my previous observations that some students may be reluctant to write in pairs and reported survey findings (e.g., Hyde, 1993), students in both classes were given the choice of working in pairs or individually. Of the 23 students present on the day of data collection, 18 chose to work in pairs and five chose to work individually.

As can be seen from Table 1, the participants\textsuperscript{2} in the study came from a range of language backgrounds. The majority were international students from Asia. The average

---

\textsuperscript{1} It should be noted that although the task was not graded, and hence of low stakes, these students are mature age students who tend to take all class tasks fairly seriously.

\textsuperscript{2} Pseudonyms are used for all participants.
age of the participants was 23. Given that the students were placed in the class on the basis of the diagnostic ESL test results, they were fairly homogeneous in terms of their L2 writing proficiency. Their scores on the writing subtest of the diagnostic test tended to range from 5 to 6 on a 9-point scale. The pairs were self selected and consisted of two male pairs, four female pairs and three male/female pairs.

Students who chose to work in pairs were given tape recorders to record their talk as they completed the task. All completed compositions were collected. Then, within 1–4 days after the class, students who worked in pairs were interviewed individually (in the researcher’s office) about the experience of writing collaboratively. The interviews were tape recorded.

2. Data analysis

Pair dialogues and interviews were transcribed verbatim. Thus, the data set included the completed compositions and transcripts of the pair talk and of the interviews. Each data source was analyzed separately. It should be noted that the composition data of two pairs had to be discarded from the data set. The data of one pair had to be discarded because the tape recorder was not activated correctly until the last few minutes of the task. In the case of the other pair, the two students had difficulties reaching agreement, and their joint text contained a number of instances in which two versions were written side by side. Given the nature of the analysis undertaken (elucidated below), it was not possible to include the text produced by this pair in the analysis.

2.1. Analysis of the compositions

The texts completed by the students were analyzed using both quantitative and qualitative measures. Quantitative measures included measures of fluency, accuracy, and complexity. Fluency was measured in terms of the total number of words. Accuracy and complexity measures were based on a count of T-units and clause analysis.

A T-unit is defined by Hunt (1996, p. 735) as “one main clause plus whatever subordinate clauses happen to be attached to or embedded within it.” This measure, despite concerns expressed by Bardovi-Harlig (1992), is the most commonly used unit of analysis of both written and oral discourse (Foster, Tonkyn, & Wigglesworth, 2000). In order to measure for complexity and accuracy, the compositions had to be analyzed for clauses, distinguishing between independent and dependent clauses. An independent clause is one which can be used on its own (Richards, Platt, & Platt, 1992); a dependent clause must be used with another clause in order to form a grammatical sentence in English. There is some disagreement among researchers as to how to code for clauses, particularly dependent clauses. In this study, following Foster et al. (2000), a dependent clause was one which contained a finite or a non-finite verb and at least one additional clause element of the following: subject, object, complement or adverbial. This definition seemed particularly suitable for this data, given that the learners often omitted the subject element in a clause (see Appendix B).

To measure accuracy, two measures were used: the proportion of error-free clauses of all clauses (EFC/C) and the number of errors per word. Admittedly, as Bardovi-Harlig & Bofman (1989) point out, such measures do not distinguish between type or severity of
errors. However, a measure of the number of errors per word does at least account for the distribution of errors in relation to the production unit (words in this case). Errors in this study included syntactical errors (e.g., errors in word order, missing elements) and morphology (e.g., verb tense, subject–verb agreement, errors in use of articles and prepositions, errors in word forms). Errors in lexis (word choice) were included only when the word used obscured meaning. All errors in spelling and punctuation were ignored (see Appendix B).

In analyzing texts it is important to consider not only grammatical accuracy but also complexity. This is because accuracy may be achieved as a result of a learner not taking any risks in their writing and relying on simple, well-controlled forms. At the same time a trade off may exist between complexity and accuracy. The more complex the sentences produced, the more likely they are to contain errors (Foster & Skehan, 1996). Complexity reflects the writer’s willingness to engage and experiment with a range of syntactic structures, moving beyond coordination to more complex structures which include subordination and embedding. One measure of complexity is the proportion of clauses to T-units (C/T). Foster and Skehan (1999), based on their previous research, conclude that this is a reliable measure, correlating well with other measures of complexity. Another measure of complexity is the proportion of dependent clauses to clauses (DC/C), which examines the degree of embedding in a text (Wolf-Quintero, Inagaki, & Kim, 1998).

In order to check for inter- and intra-rater reliability in coding, and following the advice of Polio (1997), guidelines were formulated stating clearly what constitutes a T-unit, a clause, and an error (see Appendix B). Then, a random sample of four compositions (forming approximately 30% of the entire data set) were coded by a second rater and re-coded by the researcher 2 days after the initial coding. Intra-rater reliability for T-unit and clause identification was 99%, and inter-rater reliability was 97%. Inter- and intra-rater reliability for error counts was lower (86%). Discussion between the raters resolved all disagreements. Because of the small sample size and the fact that data was collected in intact classes, the non-parametric Wilcoxon–Mann–Whitney test (Hatch & Lazaraton, 1991) was used to determine the statistical significance of all quantitative results.

A qualitative evaluation of the written text took into consideration the content and structure of the text and task fulfillment. A 5-scale global evaluation scheme was developed which took into consideration the kind of elements identified by Swales and Feak (1994) that a data commentary type text should include (see Appendix B). Inter-rater reliability on the global evaluation of five randomly selected texts was checked, and the differences between the two raters seemed small. There was complete agreement on three compositions. On the two compositions where the two raters differed, the difference was only half a band score.

2.2. Analysis of pair dialogues

In the first instance, all dialogues were examined for any distinct phases of writing: planning, writing, and revision phases. The time spent on these phases was noted. Then all
talk was segmented into episodes. An episode varied in length from a single turn to a number of turns. Each episode was coded for what the learners seemed to focus on. Seven focus areas were identified. These are set out with examples from the transcripts in Fig. 1. All episodes were timed (in minutes and seconds), and the total time spent on these aspects of writing was calculated.

<table>
<thead>
<tr>
<th>Focus area</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task clarification</td>
<td>Y: right on do we have to write it ? or just a tell</td>
</tr>
<tr>
<td>Generating Ideas (GI)</td>
<td>E: just a tell no writing...</td>
</tr>
<tr>
<td></td>
<td>Y: if we write then it can’t record it</td>
</tr>
<tr>
<td></td>
<td>(Y&amp;E, lines 5-8)</td>
</tr>
<tr>
<td>Language related episodes (LREs)</td>
<td>A: most of the Vietnamese</td>
</tr>
<tr>
<td></td>
<td>J: of the Vietnamese speak a little English?</td>
</tr>
<tr>
<td></td>
<td>A: can …</td>
</tr>
<tr>
<td></td>
<td>J: can speak a little English</td>
</tr>
<tr>
<td></td>
<td>(A&amp;J, lines 67-70)</td>
</tr>
<tr>
<td>Structure</td>
<td>O: Figure 3 shows that there were Vietnamese and Laotians participate</td>
</tr>
<tr>
<td></td>
<td>S: participating</td>
</tr>
<tr>
<td></td>
<td>O: participating</td>
</tr>
<tr>
<td></td>
<td>S: is correct?</td>
</tr>
<tr>
<td></td>
<td>O: ah … because there were we are using there were</td>
</tr>
<tr>
<td></td>
<td>(O&amp;S, lines 16-21)</td>
</tr>
<tr>
<td>Interpreting Graphic prompt</td>
<td>M: first we describe</td>
</tr>
<tr>
<td></td>
<td>C: describe the before first after that we should compare</td>
</tr>
<tr>
<td></td>
<td>M: describe after that yeah after that we talk about … we describe this</td>
</tr>
<tr>
<td></td>
<td>(C&amp;M, lines 42-46)</td>
</tr>
<tr>
<td>Reading/re-reading</td>
<td>N: the percentage … the percentage is about thirty ah … it is about forty percent</td>
</tr>
<tr>
<td></td>
<td>R: forty percent …is it?</td>
</tr>
<tr>
<td></td>
<td>N: yeah forty for advanced and forty for intermediate</td>
</tr>
<tr>
<td></td>
<td>(N&amp;R, lines 158-160)</td>
</tr>
<tr>
<td>Other</td>
<td>V: Figure 1 shows the comparison of English speaking skills for the Vietnamese and Laotians before they came to Australia and now ok…</td>
</tr>
<tr>
<td></td>
<td>(V&amp;T, lines 54-55)</td>
</tr>
<tr>
<td></td>
<td>C: perhaps you should write</td>
</tr>
<tr>
<td></td>
<td>M: yeah I write</td>
</tr>
<tr>
<td></td>
<td>(C&amp;M, lines 14-15)</td>
</tr>
</tbody>
</table>

Fig. 1. Coding of episodes.
2.3. Student interviews

Recorded student interviews (of all 18 students who chose to work in pairs) were transcribed. The students’ attitudes to pair work in general and to collaborative writing were collated and summarized.

3. Results

Before presenting the results, it should be noted that the time students took to complete the task varied. Individuals took between 10–15 min to compose the text (the exact time for each participant was not recorded), whereas pairs took on average 22 min to complete the text.

3.1. Comparing individually and jointly written texts

The following tables summarize the results of the quantitative analysis. Table 2 summarizes the results for the seven pairs. Table 3 summarizes the results for the five students who chose to work individually.

Tables 2 and 3 show that pairs tended to compose much shorter texts than students who composed individually, even though the pairs spent more time on the activity. The average length of the text composed by pairs was 112 words (S.D. = 17) whereas by individuals 137 words (S.D. = 43). However, the larger standard deviation for individuals suggests greater variation between individually composed texts.

In terms of both accuracy and complexity, texts produced by pairs seemed better than those produced by students individually. Texts composed by pairs appeared more accurate. The EFC/C for pairs averaged 52% (ranging from 33 to 70%, S.D. = 15.5%), and the mean error ratio was 0.07. In the case of individually composed texts, the EFC/C rate was 47% (ranging from 31 to 64%, S.D. = 12.2%), and the mean error rate was 0.09. Most of these errors related to verb tense choice, use of articles and prepositions, and omission of sentence elements (in the subject position).

Moreover, pairs tended to write more complex sentences, as measured by the length of the T-units in words, the ratio of clauses to T-units, and the percentage of dependent clauses. In the texts produced by pairs, the average length of the T-unit was 16 words, compared to just over 12 in the texts produced by individuals. The T-units in the texts produced by pairs were not only longer but also contained on average almost two clauses (1.86). Over half (55%) of these clauses were dependent clauses, whereas in the texts produced individually, the ratio of clauses to T-units was 1.61, and only 41% of clauses were dependent clauses.

Table 4 presents the qualitative scores given to the texts written by the students. The scores for all participants ranged from 3 to 4.5. A score between 3 and 3.5 is generally the norm in this class, given that these students are of intermediate L2 proficiency. What is

---

4 This number underestimates the time taken to complete the tasks in pairs because one pair (Howard & Sam) turned off the tape recorder when they were doing the actual writing or re-reading their composed text.
<table>
<thead>
<tr>
<th>Pair</th>
<th>Time on task (minutes)</th>
<th>No. of words</th>
<th>No. of T-units</th>
<th>No. of clauses (C)</th>
<th>Dependent clause (DC)</th>
<th>No. of errors</th>
<th>No. of error free clauses (EFC)</th>
<th>EFC/C (%)</th>
<th>Errors/words</th>
<th>Word/T</th>
<th>C/T</th>
<th>DC/C (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charley &amp; Mai</td>
<td>23:45</td>
<td>121</td>
<td>9</td>
<td>11</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>54</td>
<td>0.06</td>
<td>13.44</td>
<td>1.22</td>
<td>18</td>
</tr>
<tr>
<td>Ryan &amp; Noriko</td>
<td>27:58</td>
<td>133</td>
<td>9</td>
<td>16</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>50</td>
<td>0.08</td>
<td>14.78</td>
<td>1.78</td>
<td>50</td>
</tr>
<tr>
<td>Angela &amp; Jasmine</td>
<td>25:15</td>
<td>119</td>
<td>7</td>
<td>13</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>62</td>
<td>0.04</td>
<td>17</td>
<td>1.86</td>
<td>54</td>
</tr>
<tr>
<td>Victor &amp; Tanako</td>
<td>18:38</td>
<td>96</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>30</td>
<td>0.07</td>
<td>19.2</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>Yong &amp; Ed</td>
<td>26:22</td>
<td>82</td>
<td>4</td>
<td>10</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>70</td>
<td>0.06</td>
<td>20.5</td>
<td>2.5</td>
<td>70</td>
</tr>
<tr>
<td>Olivia &amp; Shirley</td>
<td>26:03</td>
<td>116</td>
<td>7</td>
<td>17</td>
<td>13</td>
<td>6</td>
<td>11</td>
<td>65</td>
<td>0.05</td>
<td>16.57</td>
<td>2.43</td>
<td>65</td>
</tr>
<tr>
<td>Howard &amp; Sam</td>
<td>7:54</td>
<td>116</td>
<td>10</td>
<td>18</td>
<td>9</td>
<td>13</td>
<td>6</td>
<td>33</td>
<td>0.11</td>
<td>11.6</td>
<td>1.8</td>
<td>50</td>
</tr>
<tr>
<td>Average (S.D.)</td>
<td>22:20 (6:50)</td>
<td>112 (17)</td>
<td>7.3 (2.2)</td>
<td>13.6 (3.4)</td>
<td>7.4 (3.4)</td>
<td>7.7 (3)</td>
<td>7 (2.4)</td>
<td>52 (15.5)</td>
<td>0.07 (0.16)</td>
<td>16 (3.14)</td>
<td>1.86(0.4)</td>
<td>55 (17)</td>
</tr>
</tbody>
</table>
Table 3
Quantitative analysis of texts produced by individual learners

<table>
<thead>
<tr>
<th>Student</th>
<th>No. of words</th>
<th>No. of T-units</th>
<th>No. of clauses (C)</th>
<th>Dependent clauses (DC)</th>
<th>No. of errors</th>
<th>No. of error free clauses (EFC)</th>
<th>EFC/C (%)</th>
<th>Errors/words</th>
<th>Words/T</th>
<th>C/T</th>
<th>DC/C (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayako</td>
<td>182</td>
<td>18</td>
<td>27</td>
<td>11</td>
<td>17</td>
<td>14</td>
<td>52</td>
<td>0.09</td>
<td>10.11</td>
<td>1.5</td>
<td>41</td>
</tr>
<tr>
<td>Arthur</td>
<td>181</td>
<td>16</td>
<td>19</td>
<td>5</td>
<td>12</td>
<td>8</td>
<td>42</td>
<td>0.07</td>
<td>11.31</td>
<td>1.18</td>
<td>26</td>
</tr>
<tr>
<td>Ulia</td>
<td>85</td>
<td>7</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>64</td>
<td>0.06</td>
<td>12.14</td>
<td>1.57</td>
<td>36</td>
</tr>
<tr>
<td>June</td>
<td>120</td>
<td>8</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td>4</td>
<td>31</td>
<td>0.15</td>
<td>15</td>
<td>1.63</td>
<td>38</td>
</tr>
<tr>
<td>Jeffrey</td>
<td>118</td>
<td>6</td>
<td>13</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>46</td>
<td>0.10</td>
<td>19.67</td>
<td>2.17</td>
<td>62</td>
</tr>
<tr>
<td>Average (S.D.)</td>
<td>137 (43)</td>
<td>11 (5.6)</td>
<td>17 (6.6)</td>
<td>6.6 (2.9)</td>
<td>12.8 (5.2)</td>
<td>7.8 (3.8)</td>
<td>47 (12)</td>
<td>0.09 (0.3)</td>
<td>12.47 (4)</td>
<td>1.61</td>
<td>41 (13)</td>
</tr>
</tbody>
</table>
interesting to note is that three of the texts produced by the pairs scored 4 or above, no individually completed texts scored above 3.5. Indeed, the average score for texts produced by pairs was 4.1 compared to 3.3 for individually produced texts.

A closer analysis of the texts produced revealed that individual writers tended to produce overly detailed texts. Three out of the five texts produced by the learners writing individually included far too much detail, restating the information given in the chart in words, rather than making generalizations on the basis of the information given. Swales and Feak (1994) note that one of the problems with data commentary texts is the tendency to repeat in words all the information shown in figures. The following extract from Ayako illustrates this problem. Not only did she include far too much detail, the details also do not support the thesis statement (English level was quite high):

First of all, I will compare both countries English level before came to Australia. In the case of Vietnamese, we can say that their English level was quite high, less than 10% people said their English was advanced, more than 10% was intermediate, more that 60% was low and a few was no English.

In contrast, the texts produced by the pairs tended to include less detail and contain clear highlighting statements which generalised the findings. This is illustrated in the following extract from the text produced by Angela & Jasmine:

Before they came to Australia, the level of English language fluence [sic] for Laotian people is lower than the Vietnamese’s. Sixty percent of Vietnamese could speak a bit English while the Laotian people couldn’t speak or speak a little.

Thus, students working in pairs produced shorter but more grammatically accurate and linguistically complex texts. Pairs also produced texts that had a clearer focus. However, it should be noted that the results of a Wilcoxon–Mann–Whitney test (α set at 0.05) showed that these differences were not statistically significant.

### 3.2. Process of collaborative composition

The analysis of the pair dialogues, summarized in Table 5, shows that although all the pairs spent some time on the planning phase, that phase in most cases was very brief (about 4–20 turns, taking up approximately 1 min on average). Most of the time was spent on the
writing phase. Only one pair (Yong & Ed) had a long planning phase (9 min), but that was because the pair misunderstood the instructions (they did not realise that they needed to write the text rather than just discuss it until the researcher brought this to their attention).

The planning phase was spent on generating some preliminary ideas, reading, and clarifying instructions, and deciding on who would be the scribe. In the case of four pairs, it also included a brief discussion of the graph. This seemed to create a structural framework which guided their subsequent writing.

As Table 5 shows, only three pairs engaged in a distinct revision phase, despite the explicit task instructions. Furthermore, this revision phase, like the planning phase, was very brief (just under 2 min). Only one pair (Howard & Sam) deliberated over language use only in the revision stage; all the other pairs deliberated over language use throughout the writing process.

Table 6 summarizes the time the pairs spent on the different activities involved in the task. Given the brief amount of time students spent on task clarification, these episodes were included in the final category, other. Time is shown in minutes and seconds and expressed as a percentage of the total time on the task.

Table 6 shows that the most time consuming activity was generating ideas (an average of 53% of the time, S.D. = 15%). Although all pairs spent some time on language deliberations, the amount spent varied. Some pairs spent as much time on generating ideas as on language (e.g., Charley & Mai; Olivia & Shirley) whereas other pairs spent very little time on language deliberations (e.g., Yong & Ed; Angela & Jasmine). These findings are in line with Cumming’s (1989) findings. Cumming, who used think aloud protocols to elicit what aspects of writing learners pay attention to, found that learners pay most attention to

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Time spent on the different phases of writing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total time on task (minutes)</td>
</tr>
<tr>
<td>Charley &amp; Mai</td>
<td>23:45</td>
</tr>
<tr>
<td>Ryan &amp; Noriko</td>
<td>27:58</td>
</tr>
<tr>
<td>Angela &amp; Jasmine</td>
<td>25:15</td>
</tr>
<tr>
<td>Victor &amp; Tanako</td>
<td>18:38</td>
</tr>
<tr>
<td>Yong &amp; Ed</td>
<td>26:22</td>
</tr>
<tr>
<td>Olivia &amp; Shirley</td>
<td>26:03</td>
</tr>
<tr>
<td>Howard &amp; Sam</td>
<td>7:54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Time spent on the different activities of the task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Generating ideas</td>
</tr>
<tr>
<td>Charley &amp; Mai</td>
<td>8:40 (36%)</td>
</tr>
<tr>
<td>Ryan &amp; Noriko</td>
<td>13:00 (46%)</td>
</tr>
<tr>
<td>Angela &amp; Jasmine</td>
<td>19:30 (78%)</td>
</tr>
<tr>
<td>Victor &amp; Tanako</td>
<td>10:19 (54%)</td>
</tr>
<tr>
<td>Yong &amp; Ed</td>
<td>21:00 (60%)</td>
</tr>
<tr>
<td>Olivia &amp; Shirley</td>
<td>11:43 (44%)</td>
</tr>
<tr>
<td>Howard &amp; Sam</td>
<td>4:37 (56%)</td>
</tr>
<tr>
<td>Average (S.D.)</td>
<td>53% (15)</td>
</tr>
</tbody>
</table>
generating ideas followed by attention to language. Table 6 also shows that students spent some time on dealing with the structure of the text and with interpreting the graphic prompt. The time spent on these aspects of writing was relatively brief—perhaps because of the nature of the task (data commentary).

It was difficult to establish whether the amount of time spent on language deliberations correlated to grammatical accuracy. For example, Yong & Ed, who also achieved the highest accuracy (and complexity) scores, spent the least time on language deliberations. However, the pairs Olivia & Shirley and Charley & Mai who also achieved high accuracy scores (EFC/C of 65 and 64%, respectively), spent a considerable amount of time on language deliberations (close to 40% of the time).

Variations in the approach to writing were noted among these seven pairs. Some pairs (e.g., Charley & Mai, Olivia & Shirley, Angela & Jasmine) adopted a recursive approach. That is, they generated an idea and read and re-read it to evaluate it for accuracy and expression before proceeding to generate the next idea. Others (e.g., Victor & Tanako) composed large chunks of the text or the entire text (Howard & Sam) and then evaluated the text composed. Other variations included the nature of input into the decision making process. Most of the pairs collaborated in the creation of the text by completing each other’s ideas, offering alternative suggestions, and feedback. These are illustrated in the excerpts that follow. However, in the case of one pair (Victor & Tanako), one student (Victor) took charge of the task and completed it mostly on his own. Tanako’s contribution was fairly minimal.

Excerpt 1: Generating and co-constructing ideas

Excerpt 1, from the data of Olivia & Shirley, illustrates the process of co-construction. The learners provide input to this process, building and adding to each other’s contribution (e.g., lines 248–249; 250–251) or offering alternative phrasing (line 250, 252). The resultant concluding sentence is more linguistically complex than the initial suggestion (line 249).

Excerpt 2: Language related episode

Excerpt 2 is an example of collective scaffolding. Here the learners (Charley & Mai) pool their linguistic resources to produce a more grammatically accurate text. Charley
suggests the use of the verb improve instead of the verb are (line 163), and Mai accepts the suggestion but amends it for aspect (line 164).

Excerpt 3: Generating an idea and language related episodes

<table>
<thead>
<tr>
<th>Line</th>
<th>Speaker</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>286</td>
<td>Y</td>
<td>the percentage of Laotians, Laotians</td>
</tr>
<tr>
<td>287</td>
<td>E</td>
<td>no English skills</td>
</tr>
<tr>
<td>288</td>
<td>Y</td>
<td>Laotians with no English ... English skills</td>
</tr>
<tr>
<td>289</td>
<td>E</td>
<td>skills much higher</td>
</tr>
<tr>
<td>290</td>
<td>Y</td>
<td>skills is ... was</td>
</tr>
<tr>
<td>291</td>
<td>E</td>
<td>was much higher yep</td>
</tr>
<tr>
<td>292</td>
<td>Y</td>
<td>was ... much higher</td>
</tr>
</tbody>
</table>

Excerpt 3 also shows a process of co-construction. Yong incorporates Ed’s suggestions but amends them for grammatical accuracy (e.g., lines 288, 290). Ed contributes to this co-construction not only by providing input but also perhaps by reassuring Yong in his deliberations over verb tense (line 290).

What the above excerpts show is that pair work provides the learners opportunities to co-construct texts, pool their linguistic resources (collective scaffolding), and thus compose more linguistically complex and grammatically accurate texts. Peers may also provide each other with explanations (see example of an LRE in Fig. 1) and reassurances.

3.3. Students’ attitudes to pair work and to collaborative writing

Interviews with the students in the days immediately following the class yielded some interesting insights about the students’ reactions to collaborative writing. All students were positive about group and pair work. However, although most (16) were positive about collaborative writing, two students felt that pair/group work is best relegated to oral activities, such as group discussions, rather than writing activities. Furthermore, of the 16 students who were positive about the experience of writing in pairs, five students did express some reservations about the experience.

Of those that found the experience positive, the predominant reason given (by 12 students) was that it provided them with an opportunity to compare ideas and to learn from each other different ways of expressing their ideas. For example Angela said:

Ah I think ... when I’m working in pairs we can get more ideas ... because different people have different ideas. So we can comparing the important ideas together that make a paragraph.

The students noted that such a pooling of resources provided opportunities to observe and to learn from each other. As Ed said:

I see him writing and I ... in this situation oh writing in this way is good. I learn, I learn much so ... I learn from him and maybe he learn from me.

Angela described how this collaborative process of pooling ideas took place:

I may give a ... I think, if I think I have a better idea, I can, I would tell her. And I think I’m listen what she thinks and then ... If she can give another better idea and I’ll accept her idea.
A similar process was described by Shirley:

Like sometimes she will say like um . . . start like the sentence, so I will think it’s good idea, just write down. And then if I think maybe another idea so you can . . . add to the sentence or delete.

Other reasons why collaborative writing was seen as positive were related to language. A number of students (six) noted that it was helpful for both improving their grammatical accuracy and learning vocabulary. In terms of improving grammatical accuracy, students commented that it is easier to correct other people’s errors than their own. As Maria put it:

. . . and you can’t hear your mistakes. But you can hear other people’s mistakes so I think it’s very helpful.

It also enabled students to observe and learn different words, as noted by Noriko:

I just watch vocabulary or . . . what vocabulary he was using, he used and . . . Well if he used the vocabulary which I didn’t know, I tried to use it for next time.

Finally, four students also noted that collaborative writing was a novel, fun activity. As Howard said:

This . . . quite interesting because I never had that kind of activity before. Well, that’s a pretty new idea for me, so I quite enjoy it.

However, some students had reservations about collaborative writing. One such reservation stemmed from a lack of confidence in one’s own language skills as well as concerns about criticizing others. This was expressed clearly by Maria:

Yeah it’s very hard because you can’t say . . . I mean if I say something I . . . I think . . . maybe she’s no . . . I can’t explain you . . . Maybe I think she’ll . . . she’ll think that I will . . . I want to be better than her . . . You know, you can’t just say stop you are wrong . . . or maybe, maybe I am wrong. So it’s hard to work in a group but it’s very helpful.

This concern about hurting the feelings of others was also noted in research on peer review tasks, particularly among Chinese students (see Carson & Neslon, 1996; Nelson & Carson, 1998).

Another concern was related to the view of writing as an individual activity. Yong, who saw merits in collaborative writing, still had some reservations:

I think ah . . . discussing the idea is quite interesting and useful but ah . . . Writing in pairs is . . . only . . . unbalanced. I think because the writing task is actually aimed at the individual’s ability.

The two students who did not find the experience positive were both female students from Japan. Both felt that working in pairs made it harder for them to concentrate, implying once again that writing is an individual, solitary activity. Noriko said that:

Oh . . . um . . . I really did not like to work with somebody for writing task because sometimes it’s going to be harder to concentrate with what, what I wanted to . . . write
or describe the chart or whatever. And sometimes really confused because . . . when I thought about that one . . . but the people said the difference things and I will, I don’t know which . . . which is correct or not. And sometimes I’m embarrassed to say or to give some opinion for the writing task.

Interestingly, both students felt embarrassed by their perceived poor English skills, suggesting quite strongly the fear of losing face. For example, Noriko, in response to the interviewer’s question about why she felt embarrassed to write collaboratively, replied:

Um . . . I don’t have much confident for my English skills.

Similar sentiments were expressed by Tanako:

Because I don’t know if I’m right or not, and I’m a bit feel nervous writing in front of people as well because my spelling is might be bad or . . .

4. Discussion

Although group and pair work are now used widely in the language classroom, getting students to compose in pairs is a fairly novel strategy. Hence, it is important to investigate the texts students produce when composing in pairs, how students go about composing in pairs, and their reactions to this experience.

A comparison of the products (completed texts) of pairs and individuals showed that pairs produced shorter texts, but texts that had greater grammatical accuracy and linguistic complexity, and were more succinct. They seemed to fulfill the task more competently. However, these conclusions are only suggestive. The quantitative results were not statistically significant. This is perhaps not surprising, given the small scale nature of the study and the relatively short texts the students wrote. The findings do suggest that the effects of collaboration on the product need to be investigated further with a larger sample size and longer texts.

In terms of the process of writing that students engaged in when composing in pairs, an analysis of the dialogues showed that, despite the variations in the approach adopted by the pairs, collaboration afforded the students the opportunity to interact on different aspects of writing. In particular, it encouraged students to collaborate when generating ideas. This was the activity the pairs spent most time on and most valued. As reported in the interviews, it enabled students to discover ideas together and exposed them to different views.

Collaborative writing also afforded students the opportunity to give and receive immediate feedback on language, an opportunity missing when students write individually. This may explain why pairs tended to produce texts with greater grammatical accuracy and complexity than individual writers. Some studies on peer review tasks (e.g., Nelson & Carson, 1998) have reported that students often do not view their peers’ feedback on grammar and lexis as effective and hence do not pay much attention to it in their subsequent revision (Nelson & Murphy, 1993). As McCarthey and McMahon (1992) point out, when peers are involved only in peer editing, they have no ownership of the text and thus little power to effect change. In collaborative writing, joint responsibility over the creation of the
text means that students may be more receptive to peer suggestions and feedback comments.

The students’ feedback on the experience of collaborative writing was overall very positive. However, the reservations expressed by some of the students, and the non-collaborative pattern of interaction noted in the case of some pairs (Victor & Tanako and the pair whose data was excluded from analysis) suggest that such an activity needs to be implemented with prior class preparation. Wilhelm (1999) lists a number of activities which can promote collaboration and develop positive interpersonal relations.

One of the main problems with collaborative writing (and peer review) tasks implemented in writing classes is that such activities generally aim to prepare students for the eventual individually written assignment (as was indeed the case in this study). As Reither and Vipond (1989, p. 855) point out, they tend to be introduced in classes “as overlays on courses still otherwise governed by traditional preoccupations.” To truly prepare students for collaborative writing may require a re-conceptualization of classroom teaching and assessment practices.

The difficult task facing writing teachers is how to respond to students’ preferences, including the preference to work alone, but at the same time develop the students’ flexibility (Kinsella, 1996), and prepare them for a future which may require them to write collaboratively (Ede & Lunsford, 1990; Murray, 1992).

References

Daiute, C. (1986). Do 1 and 1 make 2? Patterns of influence by collaborative authors. Written Communication, 3(3), 382–408.


Appendix A. Composition task

A.1. Instructions

In pairs, write a paragraph describing the information given in Fig. A1 below. When you finish your writing, check your text for grammatical accuracy.

Appendix B. Guidelines for coding and assessing writing

B.1. T-units

A T-unit is defined as an independent clause and all its attached or embedded dependent clauses; e.g., Fig. A1 depicts the comparison of the fluency of English/ between migrants coming Vietnam and Laotian/ when they arrive at the first time in Australia and now.//(1 T-unit, the end of which is denoted by // composed of three clauses separated by/ as shown).

Run-on sentences are counted as 2 T-units with an error in the second T-unit, e.g., the proportion of no English in Vietnamese and Laotian shows a big contrast, // less than 10% of Vietnamese were non-English speaker.//(2 T-units, each composed of one clause).

Sentence fragments (where the verb or copula is missing) is still counted as a T-unit, e.g., All Laotian ability low and no English.//

A coordinate clause with no grammatical subject is counted as a separate T-unit, e.g., and still have 39%/ who only can use limited English.//

![Fig. A1. English language fluency.](image-url)
Sentences in parentheses are counted as separate T-units.

B.2. Clauses

*Independent clause.* A grammatical structure which contains a subject and a verb and can stand on its own.

*Dependent clause.* A clause which contains a finite or a non-finite verb and at least one additional clause element of the following: subject, object, complement or adverbial. In the example of the T-unit above the first clause is coded as an independent clause, the next two as dependent clauses.

B.3. Errors

Tense/aspect/mood or verb formation errors are counted as one error; e.g., the figure divides four sections (coded as an error in verb use and an error in prepositions).

Tense/aspect are coded according to preceding discourse rather than looking at a sentence in isolation.

*Number of errors.* Coded according to minimal number of corrections required to make the clause error free, e.g., there were no Laotian English language fluency at the advanced and intermediate level (count as one error in omitted preposition ‘with English language fluency . . .’).

*Article errors.* Lenient about errors in articles with proper nouns (e.g., Laotian), but omitted article coded as an error in expression of quantity with a proper noun (e.g., half of Laotian).

*Omitted plural makers.* Coded as an error with the exception of the word Laotian. The word Laotian in the singular (instead of plural) appeared on the graphic prompt.

Count errors in word choice only when the chosen word/expression distorts meaning or is considered incontrovertibly wrong. An erroneous expression containing more than one word is still counted as one error; e.g., a few was no English (verb choice unacceptable) without no-English level left (expression ‘without no-English level’ counted as one error in expression).

Do not count errors in punctuations, capitalization or spelling.

B.4. Guidelines to global evaluation of writing

The writing is assessed on a score out of 5. This score evaluates the writing mainly in terms of structure and task fulfilment. In order to fulfil the task, the writing needs to include the kind of elements required in a data commentary type text (*Swales & Feak, 1994*); namely a topic sentence which makes reference to the graphic prompt and includes a summary type statement; generalizations highlighting the main findings. The text may also include some examples or evidence from the graph to support the generalizations. Half scores may be awarded (see footnote 2).

5. This is a very well written text. It is well structured. It contains a clear and complete topic sentence. Ideas are clearly organised and good use is made of linking words/phrases. The main findings are clearly stated and are supported by relevant examples.
4. This is a good text. The text has a clear overall structure. It has a clear topic sentence and ideas are generally well organised and easy to follow. Linking words/phrases are generally used appropriately. The main findings are clearly stated but may not always be supported by relevant examples.

3. This is a satisfactory text. It has an overall structure, but the topic sentence may be incomplete and linking words/phrases may be missing or used inappropriately. The main findings are not always clearly stated. Support offered may be overly detailed at times or unclear.

2. This is an adequate text. The text is difficult to follow because ideas are not well organised or linked. The topic sentence may be missing. The main findings are not clearly stated. Supporting evidence is overly detailed.

1. This is a poorly written text. It is poorly organised and difficult to follow. Main ideas are absent. Instead figures are re-stated in words.