A Case Study of Peer Feedback in a Chinese EFL Writing Classroom

MEI Ting & YUAN Qian
Huazhong Agricultural University

Abstract

The present study investigated peer feedback provided by 11 students in a Chinese EFL writing classroom. The objectives of the research were to understand how much peer feedback was incorporated into revisions, what kinds of revisions were made, and whether the revisions could lead to improvement in the students’ essays. Text analyses of all the 11 students’ first and second drafts were conducted by measuring accuracy, fluency, grammatical complexity, and vocabulary complexity. Results showed that the students incorporated a substantial part of the peer feedback in their revisions, most of which were surface-level revisions. The revised drafts were slightly improved in terms of fluency but greatly improved with respect to accuracy. No significant differences were found with respect to grammatical and lexical complexity. Results also indicated that peer-review activities could induce self-correction among students and cultivate independent critical readers and writers.

Key words: peer feedback; revision types; revision quality

1. Introduction

Second language writing instruction has gradually evolved from the traditional product-oriented approach to the process writing approach over the last three decades. Instead of focusing solely on formal accuracy and the final product of writing, process approach instills “greater respect for individual writers and for the writing itself” (Hyland, 2003: 17). It is a recursive procedure (Emig, 1983) which advocates pre-writing activities, encourages peer review, and requires multiple drafts. Peer review is a crucial component in multi-drafted process-oriented writing instructions. This paper reports on a study of peer review, more specifically, peer written feedback, in a Chinese EFL writing classroom.
2. Literature Review

The use of peer review has been generally supported in the literature as a “potentially valuable aid for its social, cognitive, affective, and methodological benefits” (Rollinson, 2005: 23). The beneficial impact and effectiveness of peer feedback have been substantiated by a number of empirical studies (e.g., Min, 2006; Paulus, 1999; Tsui & Maria, 2000; Villamil & de Guerrero, 1998). It has been claimed that peer readers can provide useful feedback (Rollinson, 1998), that writers can and do revise effectively on the basis of feedback from peer readers (Mendonça & Johnson, 1994), that peer feedback can be seen as complementary to teacher feedback in that it is more specific (Berg, 1999; Caulk, 1994), and that by reading others’ writing as critical readers, students could become more critical readers and revisers of their own writing (Rollinson, 2005).

However, these enthusiastic claims and positive findings of peer feedback have not gone unchallenged. Conflicting findings have been reported on its effectiveness and helpfulness. Some studies affirmed that peer review failed to induce much revision and did not lead to great improvement in writing. Leki (1990) upheld that students tend to respond to surface errors instead of semantic or textual ones. Nelson & Murphy (1993) found that L2 students tend to have a distrust of their peers’ feedback because they consider their peers no more knowledgeable than themselves in providing sensible feedback and thus do not incorporate peer feedback into their writing. In an interview conducted by Nelson & Carson (1998), students acknowledged that they incorporated teacher feedback in their revisions much more frequently than peer feedback. Some students were even reported to be “unsure of their own strength as competent readers” while reading others’ writings (Lockhart & Ng, 1993: 17). Cultural background was also presumed to be a confounding variable of peer feedback (Zhang, 1995). For example, Carson & Nelson (1996) reported that peer feedback activities were ineffective for Asian students who were used to teacher-dominated pedagogies and preferred to incorporate teacher feedback because the teacher was deemed to be the expert and the only source of authority. They also noted that Chinese students tended to maintain group harmony in peer interactions, thus being reluctant to speak up and give negative feedback about their peers’ writings.

Despite the benefits and drawbacks of peer feedback being alternately reported, the practice of peer review has been widely employed in writing classrooms over the last two decades. In China, where writing instructions in most high schools are still product and test oriented, the process writing approach has been burgeoning at the university level since the early 1990s (e.g. Cai, 1991; Yu & Zhang, 1996). However, it was not until late in 2000 that some reports of the practice and empirical research on peer review activities in monolingual classes in China began to appear (e.g. Chen, 2005; Deng et al., 2003; Fei, 2006; Mo, 2005; Miao, Richard & Yu, 2006; Zhang, 2008).

Mo (2005) conducted an experimental study to investigate the feasibility of organizing peer review activities in Chinese classrooms and the effects that peer review might have on student revision. It was concluded that peer review was as effective as relying solely on teacher feedback. The students were reported to be willing to receive
feedback from peers, and they were already capable of effective peer revision. The results were much more positive about peer review than those aforementioned studies in the Chinese context (e.g. Carson & Nelson, 1996; Zhang, 1995).

Miao, Richard & Yu (2006) compared two groups of students who received feedback from their teacher and their peers respectively, and peer feedback was found to have less impact than teacher feedback because students incorporated less peer feedback than teacher feedback in revisions, which thus led to less improvement than would have been possible. However, it was pointed out that despite its reduced impact, peer feedback did lead to improvements, encouraged learner autonomy, and served as a useful adjunct to teacher feedback even in Chinese classrooms, which were claimed to grant great authority to the teacher.

Nevertheless, the effectiveness and helpfulness of peer feedback also met challenges in China. For example, Fei’s (2006) study challenged some beliefs about the affective and pedagogical advantages of peer review. The students, who felt doubtful about the quality of peer suggestions and hesitated to use peer comments in revision, were found to have very negative perceptions of the helpfulness of peer review. Fei noted that although peer review has been widely used and adopted by many ESL writing teachers, it did not appear to be beneficial to her students. At the same time, Fei suggested that training students to do peer review is essential.

The conflicting findings of previous studies both at home and abroad added to writing teachers’ puzzles. They may “question the value of peer review within their particular context, or wonder how such a time consuming activity can be reconciled with course or examination constraints” (Rollinson, 2005: 23). To what extent does peer feedback benefit Chinese students? Should peer review activities be encouraged or excluded from writing classrooms? These questions call for further exploration and more empirical studies.

3. Theoretical Framework

According to Hyland (1998), feedback points in writing include symbols and marks in the margins, underlining of problems, and complete corrections, as well as more detailed comments and suggestions.

Before analyzing peer feedback, feedback points shall be categorized. Hyland (1998) divided feedback points into usable and unusable feedback. “Only the feedback which could actually be used in some way by the students in their revisions” was categorized as ‘usable feedback’ while those feedback points which “were evaluations, or offered positive reinforcement or a reader response” were categorized as ‘unusable feedback’ (Hyland, 1998: 262). More recently, Miao, Richard & Yu (2006) identified three types of feedback points: possible feedback (all kinds of feedback that could be offered to improve the students’ texts), usable feedback (all usable feedback provided by the teacher or peers), and used feedback (the feedback points that are employed by the students to revise their texts).

To examine the effects of peer feedback, revisions made in response to peer feedback
need to be classified first. Faigley & Witte’s (1981) taxonomy of revisions has been used by a number of L2 writing researchers (e.g., Paulus, 1999; Porte, 1996). According to Faigley & Witte (1981), there are mainly two types of revisions, namely surface changes and meaning changes. Surface changes, including formal changes or meaning-preserving changes, do not affect meaning and bring no new information to the text. Formal changes are “copyediting changes or proof-reading changes in areas such as spelling, tense, and punctuation” while meaning-preserving changes paraphrase existing concepts without altering the essential meanings (Paulus, 1999: 275). Meaning changes, including microstructure changes or macrostructure changes, affect the concepts and meaning by bringing new information to the text. Microstructure changes are simple adjustments or elaborations made to the existing text without affecting the overall gist of the text. They may “involve the use of cohesive ties, causing sentence sequence to be understood as consistent and parallel connected discourse” (Paulus, 1999: 274). When the overall direction and gist of the text is altered, the change is a macrostructure one. Faigley and Witte’s method of identifying revision types presents a detailed picture of revisions.

Another way to classify revisions is to classify them in terms of their success. Conrad & Goldstein’s (1999) taxonomy classified revisions into successful revisions and unsuccessful revisions. Successful revisions were “those solving a problem or improving upon a problem area discussed in the feedback”; unsuccessful revisions were “those that did not improve the text or that actually further weakened the text” (Conrad & Goldstein, 1999: 54). Revisions can also be classified in terms of the initiator of the revisions. The revisions induced by peer feedback are peer-initiated revisions; those that cannot not be traced back to peer feedback are self-initiated revisions. These two types of taxonomy may help understanding the efficacy of peer feedback.

To determine the effectiveness of peer feedback, most researchers turned to holistic scoring (e.g. Min, 2006; Paulus, 1999; Zhang, 2008). However, in contrast with holistic scoring, textual analyses could present us with a much clearer picture, showing us in detail whether the student writers’ second drafts gained overall improvement or not and more importantly, in which way the drafts were improved. Four indicators, namely, accuracy, fluency, grammatical complexity, and vocabulary complexity, “have been determined to be best measures of second language development in writing” (Larsen-Freeman, 2006: 597).

All the previously mentioned taxonomies of feedback points, revision types, and ways of determining the effectiveness of feedback served for their respective research objectives. However, each of them only focused on one particular aspect, e.g., either on revision types or on the effects of feedback, thus ignoring the on-going process of peer review. Considering this, we established an integrated research framework of peer feedback, covering the types of feedback, the types of revisions, and the indicators of L2 development in writing (as shown in Figure 1).

During the peer review process, it would be impossible for students to give their peers all possible feedback. Nevertheless, they may provide a certain amount of both usable and unusable feedback, some of which might be used in revision while some of which might be ignored or discarded. In response to feedback from their peers, students may revise their texts accordingly, some of which might be unsuccessful and some
successful. The successful revisions are usually either surface changes or meaning changes. Apart from making revisions according to peer feedback, students may sometimes initiate revisions by themselves, which might be triggered by self-discovery, learning from peers’ essays, or some other factors. Both peer-initiated and self-initiated revisions may lead to development in writing. Four indices, accuracy, fluency, grammatical complexity, and vocabulary complexity can measure the development in second language writing. In Figure 1, the solid line represents confirmed consequences of peer feedback, which have been substantiated by many empirical studies. The dotted lines indicate hypothetical consequences of peer feedback, which have been challenged or still call for more empirical studies.

Figure 1. An integrated research model of peer feedback in L2 writing (Adapted from Conrad & Goldstein, 1999; Faigley & Witte, 1981; Hyland, 1998; Larsen-Freeman, 2006; Miao, Richard & Yu, 2006)

4. Research Design

4.1 Research Questions
The following study investigated untrained peer feedback in a Chinese EFL writing classroom and was intended to address the following three questions:
1) To what extent did the students incorporate peer’s feedback into revisions?
2) What kinds of revisions did the student writers make in response to peer feedback?
3) Did the revisions lead to improvement in their essays? If so, in what ways were the essays improved?

4.2 Participants and Setting
This study was conducted in a course entitled “Basic Writing in English”. This course, lasting 3 semesters, was offered in a university in China for English majors in their second and third academic year. The course met once a week for 16 weeks in each semester, with each class session lasting 90 minutes. The objectives of the course were to help students fully develop their abilities in writing in English. The course began with sentence-level and then paragraph-level writing, and it ended with the production of different types of essays, including narration, description, exposition and argumentation. The classes were in the
form of lectures, brainstorming, group discussion, in-class and out-of-class writing, and peer-review activities

This study was conducted in the third semester. One of the researchers was the teacher of this course. In the class, there were 32 students, who were randomly divided into 9 groups, 3 to 4 in each group. All the writing assignments were conducted with a multi-draft process. After writing their first drafts, the students were required to read group members’ essays and to give each other written feedback that could be used for revision. After having received the feedback from their classmates, the students revised and wrote a second draft. The teacher then read the second drafts and gave students additional feedback. Finally, the students revised again and wrote their third draft. To probe into the students’ revision process, the researchers randomly chose three groups (totaling 11 students) from the nine original groups as participants in this study.

This was not the students’ first experience with peer review. In the first two semesters of the course, the teacher had been encouraging them to do peer-review for each other. Those who did well in peer-review were praised in class. The teacher sometimes showed the whole class PowerPoint slides with examples demonstrating why some feedback points were given successfully. However, no formal out-of-class training for doing peer-review had been given to the students.

The first time the teacher made peer review performance part of the final score was in the third semester. The final score of the course comprised three parts, with writing assignments accounting for 40%, peer feedback 10%, and the final examination 50%. Each writing assignment involved at least three rounds of multi-draft essay writing. The teacher rated students’ peer review performance according to the written feedback they gave to their group members. These rules were made clear to the students at the beginning of the semester.

4.3 Data Collection
The revision processes of the 11 participants were analyzed by carefully studying the first two drafts of an expository essay written during the tenth and eleventh weeks of the third semester. This essay was the fifth one written by the students in this semester. This assignment involved developing an expository essay on the topic of causes and effects of peer pressure on college students. All of the students enrolled in the course participated in discussions and drafted outlines in class before they set about composing the first draft outside of class. They were required to bring their first drafts to the next class for peer review, after which they were given one week to revise their first drafts. When the second drafts were finished, they were handed in for the teacher’s review. The 11 participants’ first drafts containing their peers’ written feedback and their second revised drafts were collected for analysis.

4.4 Data Analysis
The first step was to code all the first and second drafts to identify four things: all possible feedback points, all usable feedback provided by peers, all unusable feedback provided by peers, and all revisions or changes made between the first and second drafts by the
students. First, the two researchers independently identified and then double-checked all possible feedback points in the students’ first drafts. Second, they identified and counted all the feedback points, including usable, unusable and used ones.

The second step was to classify the revisions made in all the second drafts. First, the researchers located each revision made in the second drafts by comparing the first and second drafts. The peer-initiated revisions were marked as “PI”; the self-initiated revisions as “SI”. The successful revisions were marked as “SR”; the unsuccessful revisions as “UR”. Second, they studied the successful revisions to determine whether they were surface changes or meaning changes.

The last step was to conduct textual analyses by measuring the four indices of accuracy, fluency, grammatical complexity, and vocabulary complexity of the first and second drafts of the 11 students. Each essay was segmented into T-units. A T-unit is defined as “a main clause plus all the subordinate clauses attached to or embedded in it” (Hunt, 1965: 141). According to Wolfe-Quintero, Inagaki & Kim (1998), accuracy was measured by the proportion of error-free T-units to T-units (EFT/T) and the numbers of errors per T-unit (E/T); fluency was measured by the average number of words per T-unit (W/T), the average number of words per clause (W/C), and the average number of words per error-free T-unit (W/EFT); grammatical complexity was measured by the average number of clauses per T-unit (C/T) and dependent clauses per clause (DC/C). Lexical complexity was measured by the type-token ratio, i.e., the number of word types divided by the square root of two times the number of words, which was calculated with the help of RANGE, a software program specialized for analyzing word types.

5. Results and Discussion

5.1 Incorporated Feedback
The students incorporated a large part of peer feedback into the revision. The total number of possible feedback points was 477. Altogether, the 11 students had 337 usable feedback points of the usable feedback points, 85.5% (288) were incorporated into their revised second drafts. Of all the 340 revisions made to the first drafts of the essays, a total of 84.7% (288) resulted from peer feedback. Most of the revisions initiated by the peers, 266 out of the 288 (92.4%), were successful. These results are summarized in Table 1.

Of the 11 students, 9 initiated 52 (15.3%) revisions by themselves. Most of these revisions, 49 out of the 52 (94.2%), were successful. This revealed a tendency for self-correction in the peer feedback group, a tendency that is in line with a finding of Miao, Richard & Yu’s study (2006): when there was less dependence on the “authoritative” teacher, some students tended to take the initiative of doing self-revision. Peer review as shown in this study seems to promote L2 writers’ autonomy, and this “has been acknowledged by many L2 writing researchers” (Suzuki, 2008: 211). Meanwhile, it also appears that by reading peers’ essays as critical readers, some students became more critical readers and revisers of their own writing.
Table 1. Amount of feedback points and revisions

<table>
<thead>
<tr>
<th>Name</th>
<th>Possible</th>
<th>Usable</th>
<th>Unusable</th>
<th>Used</th>
<th>Peer-initiated</th>
<th>Self-initiated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hai</td>
<td>39</td>
<td>30</td>
<td>4</td>
<td>19</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td>Chen</td>
<td>43</td>
<td>38</td>
<td>1</td>
<td>32</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>Kai</td>
<td>61</td>
<td>48</td>
<td>1</td>
<td>36</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>Xian</td>
<td>41</td>
<td>33</td>
<td>0</td>
<td>31</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>Sun</td>
<td>36</td>
<td>16</td>
<td>6</td>
<td>13</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Li</td>
<td>38</td>
<td>22</td>
<td>6</td>
<td>14</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Xiao</td>
<td>38</td>
<td>24</td>
<td>2</td>
<td>19</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>Xue</td>
<td>25</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Xiu</td>
<td>41</td>
<td>23</td>
<td>3</td>
<td>26</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>Dan</td>
<td>52</td>
<td>42</td>
<td>2</td>
<td>42</td>
<td>41</td>
<td>5</td>
</tr>
<tr>
<td>Hong</td>
<td>63</td>
<td>50</td>
<td>5</td>
<td>48</td>
<td>46</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>477</td>
<td>337</td>
<td>40</td>
<td>288</td>
<td>266</td>
<td>49</td>
</tr>
</tbody>
</table>

Possible: all the possible feedback points identified by two researchers; usable: All the usable feedback points provided by peers; unusable: All the unusable feedback points provided by peers; used: feedback used by students in response to feedback

5.2 Types of Revisions

Altogether, the 11 students made 340 revisions to their essays, and 82.4% (280) of them were categorized as surface changes. Among these surface changes, 37.9% (129) were formal changes and 44.4% (151) were meaning-preserving changes. Apart from the surface changes, 60 changes (17.6%) were meaning changes, categorized as either microstructure changes, accounting for 16.7%, or macrostructure changes, accounting for 0.8%. See Table 2.

Of the 288 revisions that were made as a result of peer-initiated feedback, 89.9% (259) were surface changes, whereas meaning changes only accounted for 10.1% (29). These results may show that surface changes predominated in the revisions. The students focused mainly on surface-level aspects when doing their revisions. This finding resonates with Paulus’s (1999) result that surface changes occupied much higher percentages than meaning changes did, but it is contrary to Miao, Richard & Yu’s (2006) claim that due to the students’ perception of their low linguistic abilities, peer-initiated revisions were concerned less with surface changes. This three-semester writing course began with sentence-level and paragraph-level writing. The students might have gained more chances to develop their linguistic ability and build confidence in themselves through many rounds of writing and revising. At the same time, the encouraging environment established in the classroom induced much feedback, although most were of surface-level.

Among the 52 self-initiated revisions, microstructure changes (55.8%) ranked first, followed by meaning-preserving changes (23.1%) and formal changes (17.3%). This indicated that when doing self-correction, these students focused more attention on meaning-level aspects. The 11 students varied greatly in their revisions. Four students barely made any self-initiated revisions, while the other 7 students made self-revisions. More encouraging is that these students made mainly meaning changes when doing self-revisions. The reasons why these students behaved differently in peer review activities could be studied in future research.
Table 2. Types of revisions from draft 1 to draft 2

<table>
<thead>
<tr>
<th>Name</th>
<th>Surface Changes</th>
<th>Meaning Changes</th>
<th>Microstructure Changes</th>
<th>Macrostructure Changes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Formal Changes</td>
<td>P-I</td>
<td>S-I</td>
<td>P-I</td>
<td>S-I</td>
</tr>
<tr>
<td>Hai</td>
<td>10</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chen</td>
<td>13</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Kai</td>
<td>12</td>
<td>0</td>
<td>20</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Xian</td>
<td>14</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Sun</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>10</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Xiao</td>
<td>3</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Xue</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Xiu</td>
<td>24</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Dan</td>
<td>14</td>
<td>2</td>
<td>25</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hong</td>
<td>11</td>
<td>3</td>
<td>27</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>120</td>
<td>9</td>
<td>139</td>
<td>12</td>
</tr>
</tbody>
</table>

P-I: peer-initiated changes; S-I: self-initiated changes

5.3 Revision Quality

Due to the small sample of this study, a Wilcoxon matched-pairs signed-ranks test was performed to determine if there was any statistically significant differences between the first and second drafts with respect to fluency, accuracy, grammatical complexity and lexical complexity. The alpha level was set at 0.05. The results indicate that the revisions led to improvement in the students’ essays. See Table 3.

Among the three indices of fluency, W/T, W/C, and W/EFT, only the difference between the two drafts on W/EFT was significant with the difference favoring the second drafts (z = -2.934, p < .05). This indicates that the average number of words per error-free T-unit increased in the second drafts. The significant difference of W/EFT between the two drafts echoes the findings in the first research question, that is, 82.4% of total revisions were surface changes, 37.9% of which were formal changes, focusing principally on grammatical errors and mechanics. It suggests that grammar revisions mainly led to the increase of W/EFT in the second drafts.

For the two indices of accuracy, significant differences were detected in both EFT/T (z = -2.934, p < .05) and E/T (z = -2.934, p < .05) between the two drafts. The ratio of error-free T-units to total T-units (EFT/T) increased in the second drafts while the average number of errors per T-unit (E/T) decreased. This suggests that peer feedback helped improve accuracy in the students’ essays.

There were no statistically significant differences found between the two drafts on either grammatical complexity or lexical complexity.

Overall, as grammar and mechanics were the most frequently revised aspects, the students’ essays showed slight improvement in fluency, significant improvement in accuracy, while no distinct improvement was found in terms of grammatical complexity and lexical complexity.
Table 3. Language development from the first draft to the second draft

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Index</th>
<th>Draft 1</th>
<th>Draft 2</th>
<th>Total</th>
<th>Mean</th>
<th>SD</th>
<th>z</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W/EFT</td>
<td>167.057</td>
<td>185.768</td>
<td>176.413</td>
<td>15.187</td>
<td>16.888</td>
<td>4.246</td>
<td>.003*</td>
</tr>
<tr>
<td>Accuracy</td>
<td>EFT/T</td>
<td>5.599</td>
<td>9.097</td>
<td>7.398</td>
<td>.509</td>
<td>.827</td>
<td>.122</td>
<td>.003*</td>
</tr>
<tr>
<td></td>
<td>E/T</td>
<td>9.383</td>
<td>2.673</td>
<td>5.930</td>
<td>.853</td>
<td>.243</td>
<td>.331</td>
<td>.003*</td>
</tr>
<tr>
<td>Complexity</td>
<td>DC/C</td>
<td>2.794</td>
<td>2.937</td>
<td>2.864</td>
<td>.254</td>
<td>.267</td>
<td>.080</td>
<td>.202</td>
</tr>
<tr>
<td>Lexical</td>
<td>Type-Token</td>
<td>4.851</td>
<td>4.829</td>
<td>4.838</td>
<td>.441</td>
<td>.439</td>
<td>.044</td>
<td>.635</td>
</tr>
</tbody>
</table>

1. W/T: the average number of words per T-unit; W/C: the average number of words per clause; W/EFT: average number of words per error-free T-unit; EFT/T: The ratio of error-free T-units to total T-units; E/T: the average number of errors per T-unit; C/T: average number of clauses per T-unit; DC/C: the ratio of dependent clauses to total clauses
2. (a) Based on positive ranks
3. *p < .05 (Wilcoxon Matched-Pairs Signed-Ranks test)

6. Conclusion

This study explored peer written feedback in a Chinese EFL writing class. In the study, the participants incorporated a large part of peer feedback into their revisions. The peer-initiated revisions were found to mainly focus on surface-level aspects, which led to slight improvement in fluency, significant improvement in accuracy, but no significant improvement in grammatical or lexical complexity. An encouraging finding that deserves attention is that some participants did self-revisions, most of which were meaning changes.

Although most of the revisions were of surface-level, the efforts that the students made in the revising process still represents “a cognitive and linguistic effort that needs to be acknowledged” (Villamil & De Guerrero, 1998: 505). It should also be noted that though peer review activities can not guarantee the overall improvement in a piece of writing, students can become more critical readers and revisers through reading others’ writings critically. Learner autonomy could also be a by-product of peer review activities. Peer feedback may bring benefits to Chinese EFL; therefore, peer review activities should be encouraged in writing classrooms.
This study was a small-scale study, and the results obtained cannot be generalized to other educational contexts. Furthermore, the participants in this study were seniors majoring in English, who had received more systematic training in English writing than juniors or non-English majors. Although accuracy and fluency were found to improve with the help of peer feedback, the findings might be different in other research contexts.

Nonetheless, some pedagogical suggestions may shed light on ESL writing instruction. In this three-semester writing course, although peer review was not formally trained, the teacher had always been demonstrating and encouraging peer-review. Unsurprisingly, the students complained and felt frustrated in the first semester, the same as Fei reported in her study (2006), but from the second semester onward, most of them gradually began to realize the benefits of peer-review. Therefore, peer review training, be it formal or informal, may decide the effectiveness of peer review activities. As writing teachers, we need to realize that peer-review is an on-going process. It takes time and efforts to establish an environment that encourages peer-review, which cultivates independent critical readers and writers.

Further research can be done to determine why some peer feedback points were not incorporated in revisions and why the self-corrections were mainly meaning changes. Retrospective interviews could be organized to find out about the mental process of revision, for example, how the students made decisions as to using or discarding certain peer feedback, what prompted them to make self-corrections in addition to making peer-initiated revisions.

References


(Copy editing: Julia Hofmann)