

The Effectiveness of Peer Feedback

– A Study of Collaborative Learning Techniques for Improving Writing in University English Classes

Mark Holst
Beverley Horne

University is the time for students to express their ideas in English to say what they want to say without exam pressure being their greatest motivation. Writing is not just an academic exercise for the writer, but it is meant to be read, which gives students an opportunity to share their writing with their peers and develop confidence which comes through sharing. In reader-writer and product-process relationships the first draft is improved in a more collaborative and supportive way. Therefore, university writing classes can be a time to experiment and understand why particular structures or usages are necessary or more appropriate. In order to create meaningful content students need to take ownership of their own writing.

The issue of feedback has always concerned teachers of L2 writing. While providing written feedback on assignments is an essential part of teaching writing, it can also be time-consuming. Although it is generally agreed to be beneficial to students, it can be difficult for teachers to know how much the students are learning from the feedback. In order for feedback to benefit the students, they need to understand it and engage with it actively. To facilitate this, teachers need to consider what kind of feedback is necessary, effective and appropriate. In recent years, attitudes to written feedback have also been influenced by the proliferation of online language checking and software tools. Despite these changes, one goal of L2 writing should be for students to understand how language is used to create meaningful content. To do this successfully, they need to understand their own errors in writing. This might better be achieved if they are actively involved in the process of correcting their assignments. This study investigates whether a learner-centred and task-based approach using peer collaboration to solve errors highlighted by the teacher might be an effective way to approach error correction in student writing.

In this approach, learners collaborate with classmates to consider grammatical, lexical, punctuational and discourse errors which the teacher has highlighted on their respective essays. The novelty of this approach is that it turns error correction into a communicative activity: a kind of puzzle-solving exercise where language problems are discussed in an objective way using the students' combined language resources. Each learner improves their own English accuracy and fluency identifying and correcting the highlights while benefiting from the pooled knowledge of their peers. Furthermore, as they are receiving the help of a peer they have a responsibility to take the error correction task seriously, reviewing their writing more thoroughly than they may otherwise have done on their own. Collaboration therefore takes the

pressure off learners to solve their own errors making the process more enjoyable and more fulfilling.

Research Aims

- To promote independent learning by making students take responsibility for their own English language development within a structured framework through discussions with peers after initial highlights made by the teacher.
- To use a communicative approach based on discrete tasks and identified outcomes.
- To promote students' motivation to improve their writing knowing that their essay will be looked at in detail by the teacher and their classmates.
- To promote students' motivation to boost learner confidence though noticing similar errors made by other learners.
- To enhance students' teaching skills through explaining their partners' errors

Based on these aims this investigation attempted to determine how effective the system was through considering the following points:

- Did the students collaborate effectively with their peers?
- Did the students correct the highlights through peer discussions?
- Did the students deepen their understanding of English usage through the activity?
- Did individual students improve over time?
- How did the students evaluate the system?

2. Peer Feedback and Error Correction

The need for peer feedback in writing development

In many cases, Japanese students have not had the opportunity to do much extended writing in English prior to starting university. The emphasis in secondary education has traditionally been at the sentence level with a focus on form over function. While, for many students, this may have resulted in a good declarative knowledge of English, an examination of their writing can reveal the extent of their procedural knowledge. However, moves to encourage self-responsibility in Japanese learners of English can be seen by the development of error correction resources available for self-study including the online corpus-based system (SCoRE) using an inductive based data driven learning approach to error correction (Chujo et al, 2015) and reference materials such as Barker (2010) which explains common English errors by Japanese learners.

Students have generally been educated in a traditional system where the teacher is seen as the authority, so learners are less likely to trust their peers to correct their English, (Elwood &

Bode, 2014; Sellick & Bury, 2018) so as well as the obvious pedagogical need, students expect the teacher to edit and comment on their writing. The question is at what stage of the writing process should this be done and what kind of feedback should the teacher be giving?

Bitchener (2019) examines the role corrective feedback (CF) plays in the learning process from an SLA perspective: before any learning takes place learners need to be motivated to receive feedback and be oriented towards accuracy and meaning. Then they are ready to go through Gass's (1997) five processing stages: (1) attend to the input (teacher's feedback); (2) notice and understand what the feedback is telling them; (3) analyse and compare the feedback with the original output; (4) make hypotheses about what they believe to be accurate output; (5) produce new output. In this framework stages 1~4 are highly suitable to be carried out using peer discussion, after which students redraft the original writing on their own (stage 5).

During peer discussions learners evaluate the teacher's feedback on each other's output and give comments on each other's drafts. This is thought to have a number of potential benefits for L2 writing development. There are various types of student feedback summed up as follows:

- **Written corrections of errors and offering up corrections.** Potential problems are that learners might make false corrections and that students might not trust their peers' corrections.
- **Verbal feedback:** negotiating with peers about a potential error giving both the chance to discuss it and work on solutions together, while asking the teacher about unresolved issues. However, this can be time-consuming taking up class time to address even a few errors.
- **Comments written at the end of the essay about the content and English readability.** Students are less likely to offer comments about perceived language problems because of lack of confidence in their own English and relying on the teacher to do this.

It is often claimed that teacher corrections leave nothing for the learner to do. (Carless, 2020) Despite the huge time investment in writing feedback comments, teachers cannot know with certainty to what extent the students read, understand and learn from them. It is often the case that students passively accept the corrections without asking follow-up questions. Consequently, there is no active engagement by the learner in thinking about their errors and trying to come up with a solution. Highlighting, returning the assignment to the learner but not asking for a second draft will similarly be wasted effort by the teacher. However, teacher highlighting plus peer discussion will validate the teacher's effort as the learner is required to address their mistakes with a fellow learner.

The purpose of a highlighting system is to promote language development through a

deeper consideration of structure, expanding lexical knowledge through considering the appropriateness and nuances of vocabulary and developing a deeper knowledge of discourse structure. Highlighting therefore provides a platform for mutual learning through peer discussion of error repair. However, while collaborative peer feedback is often seen as beneficial by learners, it is not trusted as much as feedback from a teacher, and therefore diminishes its potential benefits. One solution to this problem is collaborative writing: the co-authoring of a text, 'a joint-writing activity, undertaken by two or more authors, which culminates in the production of one text'. (Storch, 2019) Through this process learners are in constant negotiation, giving each other feedback and building on each other's contributions. Underpinning this approach is Vygotsky's sociocultural theory of the mind, in which all cognitive development is seen as:

occurring in social interaction, where the novices were provided carefully calibrated assistance by an expert member of the community or peer; assistance that is responsive to the novice's needs. Providing such assistance requires the active involvement of both the novice and the expert, and hence the centrality of interaction of development. (Storch 2019: 144)

Storch (2019) writes that successful collaborative feedback needs to be carried out under specific guidelines and that learners need to be trained how to do it. Learners give feedback and negotiate with each other at all stages of the writing process: how to interpret the task requirement; how to structure and link their ideas together (discourse); feedback on the language (grammatical accuracy and lexical choice); and feedback on mechanics (orthography and punctuation). Sellick (2018) found that the benefits of peer feedback are that it improves student motivation by improving student ownership of work, increasing student autonomy and developing students' ability to make judgments.

Er et al (2020; 2021) note that feedback often fails because it is a one-way transmission of diagnostic information where students play a passive role as information receivers. Peer discussion can help students make sense of teacher feedback and act on it, and it is especially useful in large classrooms if it is done systematically. They present a theoretical framework as a three-phase collaborative activity: (1) planning and coordination of feedback activities (involving socially shared regulation); (2), feedback discussion to support its uptake (involving co-regulation), and (3) translation of feedback into task engagement and progress (involving self-regulation).

Japanese university context

Colpitts and Howard (2018) carried out a study into students' response to the collaborative feedback process in a Japanese university. They found that most students valued and benefited from the process, particularly in terms of noticing their mistakes and the analytical engagement

with language which Swain (2006) defined as “*linguaging*”.

Elwood (2014) distinguishes between direct feedback and indirect feedback. Indirect feedback may engage students’ thinking skills more. Overall, he found a strong preference for direct, written feedback. He concluded that students were generally positive regarding feedback. Wakabayashi (2008) emphasizes the dual benefit of peer feedback: that learners not only receive peer feedback from others but also act as providers of feedback. She saw the value of peer feedback as a collaborative learning process which increased learner awareness. Other benefits were that the students received more feedback than they would from the teacher alone, received comments from learners’ perspectives and gained audience awareness. They could also enhance their critical thinking skills.

3. Research Method

To address the research questions a longitudinal study was carried out following the principles of action research, which we describe here.

Setting

The research was conducted in two first year classes at two Japanese universities hereafter, Group A and Group B. Group A was an English writing class on an English language teacher training course in the Faculty of Education of a large multi-faculty public university (N = 21). The students were mainly English majors in their first year (18 students) and English minor students in their second or third year (3 students). All classes were conducted face to face on campus. Group B was an English communication class as part of a general studies programme at a single faculty public university (N = 32). The advantage of having two distinct settings allowed us to compare the effectiveness of this technique with both English language majors and non-English language majors. Each group had a compulsory 90-minute weekly English class over a 15-week semester. All of the classes were face to face except for three of Group B’s classes.

Implementation

During the 15-week course students in both groups submitted six writing assignments designated by the researchers. (Appendix 1) The essays were of various discourse genres in order to practice different grammatical features (especially tense variation) and stylistic features (reporting, explanatory, giving opinions, descriptive and persuasive). Students had a week to write an essay and upload it on the university learning management system (LMS). For each essay students stated how long it took them to write it. In the following class students read and commented on each other’s essays. The teacher edited the essays using a highlighting system (Appendix 2) where different colours represented different types of mistakes (grammatical, lexical, punctuation, spelling and discourse). In the third class students collaborated in pairs, discussing how to repair the highlighted errors. Finally, students submitted a second draft, incorporating the repairs decided on with their partners and by themselves. After three of the

peer discussions students filled out online questionnaires about their experiences of the peer discussions and rewriting process, giving examples of language points they had discussed and assessing how useful or effective they found the approach. (Appendix 3) After the final peer discussion students filled in a final, more detailed questionnaire about the process as a whole. (Appendix 4)

In Group A the activity ran for 30 to 40 minutes each time, whereas in Group B it only ran for about 20 to 30 minutes. This is because Group A was a designated writing class, whereas Group B was a communication skills class in which writing was only one component of the whole.

Data collection and analysis

By the end of the course the following information had been collected:

- First drafts with highlights and second drafts of the six essays written by each student.
- Four questionnaires about the system submitted by all students.

The essays gave an external (teacher's) view of each student's writing performance while the questionnaires gave an internal (student's) view of their own writing and the error correction process. Essays were examined to see how many and what kind of highlights each student received in their first drafts, and compared with the second drafts to see how they had edited the highlights. Also, using the self-reported writing times and the word counts, simple calculations were made for error frequency (highlights/number of words) and writing speed (word count/writing time) in order to compare weaker and stronger students within each class, and to compare individual students' progress over the six essays. Questionnaire responses were analysed quantitatively (binary and multiple-choice questions) and qualitatively, categorising student comments into common themes regarding (i) the highlighted errors the students had discussed with peers and (ii) their opinions about the system itself.

4. Results and analysis: Patterns in student essays

Group A had more highlights denoting style problems than Group B, and even in Group A they were mainly accounted for by a small number of very high level students.

Table 1: Differences across the 6 essay assignments by group

	Essay length (words)			Error rate (words /highlights)			Writing Speed (words/min)		
	Mean	Hi	Lo	Mean	Hi	Lo	Mean	Hi	Lo
Group A	307	254	404	33	14.8	129*	4.0	6.9	2.0
Group B	229	183	297	19	9.2	38.3	3.0	4.5	1.3

Error Rate: low number = high error rate/ high number = low error rate. (e.g. 33 = one error every 33 words)

*A3 was an outlier – the next lowest rate (A) was 71.9

Individual differences

There was variation in: (i) the English proficiency between learners; (ii) the total number of highlights (see Table 1: Error Frequency); (iii) the success rate of highlight repair; (iv) the average length of time taken to write the essay. Some learners could repair almost all their errors (B1) while others could consistently correct a particular type of error but not a different type of error (e.g. B2, who had many mistakes in general).

Effectiveness of the highlighting system

How successful was the highlighting system in directing the students to the cause of the problem?

Some learners were able to correct a particular type of error sometimes but not other times. This may be related to ambiguity or inconsistency in the teacher's use of the highlights: when more than one or two words are highlighted (up to clause level) then the learner has less information available and may re-engineer the clause randomly or without careful thought. Some examples of highlighted sentences and corresponding inappropriate repairs are listed below.

- Misunderstanding the category of highlight (e.g. making a vocabulary repair instead of a grammar repair)
- Not highlighting all the relevant parts of the error to make sure the learner is properly focused. (e.g. when indicating infinitive vs gerund, the teacher should highlight 'to' + 'verb', not just 'to' (see)).
- Where two different highlights are side by side – the learner conflated these into one:
 - (i) 'And at [] English conversation class'
Student repair => *And in English conversation class*
Target => *And in the English conversation class*
 - (ii) 'That is [] important experience of my high school life'
Student repair => *That was important experience of my high school life.*
Target => *That was an important experience of my high school life)*

4.2 Students' development as seen through their feedback.

Language problems emerging from the highlights

Given the different profiles and settings of the two groups and the fact that the respective students' essays were marked by different teachers it was expected that there would be differences in the number of highlights between the groups as a whole. In total, Group A averaged about one highlight every 26 words while Group B averaged one highlight every 18 words (Table 2).

Table 2: Frequency of highlights per essay (= Number of words/Number of highlights)

Essay No.*	1	2	3	4	5	6	Overall
Group A	23.5	22.8	24.7	29.2	32.8	22.9	25.98
Group B	15.4	17.3	16.2	22.3	13.8	24.4	18.2

* See Appendix 1

Within each group further differences were also observed. In Group A the frequency of highlights was similar for essays 1, 2, 3 and 6 (range = 22.8 to 24.7) with essays 4 (Describe a famous person) and 5 (Explain a Japanese dish) having a higher but similar frequency (29.2 and 32.8 respectively). The high numbers may have been due to the type of assignment, both of which involved research. Meanwhile, in Group B there was a general trend for the number of highlights to decrease during the course of the six essays, from one highlight every 15 words in Essay 1 to one in 24 words in Essay 6. However, this trend was bucked by Essay 5 (Explain a Japanese dish) with one highlight every 14 words. The high number may have been due to the type of assignment, which involved research about the origins of the dish, a description of how to make it and comments justifying their choice. Group B was most successful on the final essay, which could indicate an improvement in accuracy over the course, or it could be accounted for by the type of essay which they probably had most experience of during their preparation for university entrance exams (persuasive - arguing for a position).

Types of errors and degree of successful repair.

Table 3: Language errors discussed during peer collaboration

Vocabulary	120
Missing Words	52
Spelling / Punctuation	24
Style / Discourse	66
Grammar	197
Articles	66
Verb problems (Tense/Mood)	63
Prepositions	42
Plurality & S-V Agreement	27
Sentence SVO	10
Pronouns (Inc. Relative)	7
Morphology POS	6
Others*	9

*infinitive v gerund, conjunctions, missing object, negative forms, there is/ are

Students were able to successfully repair such features as replacing nouns with pronouns, replacing singular pronouns with plurals and replacing infinitives with gerunds. They had most difficulty in addressing style highlights.

Unsuccessful repairs

On the other hand, there were a number of highlights that many students had difficulty solving: Grammar

- article mistakes,
- inability to select the appropriate preposition or collocate the appropriate preposition with a particular verb
- tenses: present used instead of present perfect: present used instead of past.
- relative pronouns: unable to replace ‘that’ with ‘when’, ‘where’, ‘which’
- sentence/clause level: difficulty in repairing highlighted phrases or clauses.

There were a few vocabulary mistakes such as using ‘the’ instead of ‘my’, or ‘then’ instead of ‘so’. There was also a tendency to assume missing word spaces should be filled by an article or determiner.

Individual students’ repair development

To further investigate the general trends, six essays from three students in Group B were examined to determine what kind of errors were being highlighted and how successful the students were in repairing them in their second drafts. The students were selected to represent the range of English ability in the group based on their total number of highlights (the top students had the fewest highlights).

Student B12

B12 was an accurate writer with only 37 highlights over the 6 essays (with only 1 highlight on essay 6), at a rate of one highlight every 36.4 words (3rd lowest in the class). B12’s overall writing speed (self-reported) was about 3.8 words per minute, well above the class average (2.9). Their average weekly score was 10 points above the class average, and their essays were all comfortably over the required length each week.

The student managed to correct 31 of their 37 highlights, the most common mistakes being with articles, but there were also problems in selecting appropriate tense forms (e.g. past perfect instead of simple past). The student was able to repair almost all of these errors, although there were some problems in selecting the appropriate article, and they were unable to replace an object pronoun with a possessive pronoun (‘A student of him’). There were also a couple of stylistic problems connected with end focus which this student was unable to repair, even though there was an example of this on the ‘Highlight Symbols’ handout.

On the feedback forms the student reported discussing the following errors with their partner: tenses, pronouns, prepositions; subject verb agreement (*e.g. He play the soccer*); word meaning; missing articles (*e.g. [] person who ...*); paragraphing. Many of these errors did not appear on the student's own writing, so it appears that they were their partner's problems and, given their own English proficiency, they may often have been engaging in more of an expert or teaching role.

Feedback Responses

This student found the feedback sessions 'Very useful' and that they were 'A little useful' for their English. They stated a preference for the multi-coloured highlighting system and for working on errors with a partner rather than by themselves. They also said there was a difference between working online or face to face. Each week they worked on more than five errors with their partner in class, all of which they were able to solve and they also worked on more errors by themselves after class, but never with another student. The main resource they used to solve their errors was a dictionary.

Student B2

B2 performed in the mid-range of Group B, having 81 highlights over the 6 essays (average = 77), at a rate of one highlight every 17.3 words (average = 20). B2's overall writing speed (self-reported) was about 1.9 words per minute, well below the class average (2.9). Their average weekly score was 10 points above the class average (second highest), and their essays were comfortably over the required length each week (Av = +26).

The student managed to correct 53 of their 81 highlights. Among the highlights they were unable to repair, two recurring problems were confusing definite and indefinite articles, and not being able to change present tenses to past tenses ('is' => 'was', 'withdraws' => 'withdrew'). On the other hand, they were successful in repairing pronoun forms (me => my, him => his, them => their), correcting singular forms to plurals, as well as a lot of success repairing highlighted clauses:

- For example, putting the humidifier or the aroma diffuser.
=> For example, I put the humidifier and the aroma diffuser there.
- Furthermore, it is not only about foreign country but also about our country to develop an understanding.
=> Furthermore, I can develop an understanding not only about foreign country, but also my country.

By the final essay this student was able to correct almost all of the highlights (11 out of 14)

The student reported discussing the following problems with their partner: fundamental grammar (SVO), 'either' versus 'too' in negative sentences (*When I write a negative sentence, I have to use not "too" but "either"*), missing articles and the difference of "the" or "a",

forgetting to put “s”, and confusing the meaning of close synonyms.

Feedback Responses

This student found the feedback sessions alternately ‘a little useful and very useful’ and that they were ‘Very useful’ for their English. They preferred the multi-coloured highlighting system and they preferred to work on their errors with a partner rather than by themselves. They also said there was a difference between working online and face to face. Each week they worked on more than five errors with their partner in class, all of which they were able to solve, and they also worked on more errors by themselves after class, but never with another student. The main resource they used to solve their errors was their partner’s comments.

Student B19

B19 received 107 highlights over the 6 essays, at a rate of one highlight every 12 words, well below the class average of 20. B19’s overall writing speed (self-reported) was about 2.3 words per minute, below the class average (2.9). Their average weekly score was 5 points below the class average, and their essays were all just over the required length each week.

The student managed to correct 73 out of 107 highlights. Of the highlights they were unable to repair notable problems were plural forms, repairing morphological mistakes (e.g. rectangle => rectangular) and prepositions: some wrong prepositions were replaced with other wrong ones and some missing prepositions were left missing. On the other hand, they were successful at filling in missing articles and replacing nouns with pronouns, evident especially in the final essay. They also managed to fill in missing auxiliary verbs (have and will). B19 also successfully dealt with a highlighted clause by reconstructing the sentence in a much simpler way, but still managing to express their main idea:

- *I am disappointed about chances to wear clothes is not many for me because of this pandemic.*

=> Because of this pandemic, I can't go out.

By Essay 6, even though they had 30 highlights (21 corrected successfully) they did a better job of correcting mistakes (especially verb tenses and articles) compared to Essay 1.

In the feedback forms this student reported discussing and solving the following errors with their partner: irregular pluralization (men → man), tenses, using Japanese English and missing articles (*the*). They also mistakenly categorized a vocabulary mistake as a grammar mistake (*to talk English* → *to communicate*), and a punctuation error as a style problem (*japanese* → *Japanese*) which suggests a need for more training in error categorization.

Feedback Responses

This student found the feedback sessions alternately ‘a little useful and very useful’ and that they were ‘Very useful’ for their English. They preferred a single-coloured highlighting system and they preferred to work on their errors with a partner rather than by themselves. They said there was no difference in working online or face to face. Each week they worked on more than five errors with their partner, all of which they were able to solve, and they usually worked on more errors by themselves after class, but never with another student. The main resource they used to solve their errors was their partner’s comments.

In this section we have focused on analysis of the students’ written work both collectively, to see general patterns within and between groups, and individually by focusing in on 3 case studies to better understand the writing and rewriting process the students underwent. In the next section we will discuss the results of the surveys.

5. Results and analysis: Student Feedback – Assessing the peer discussion system

All 52 students were surveyed after three of the peer feedback sessions (Group A - after essays 1, 5 & 6 and Group B after essays 1, 2 & 4), and both groups were surveyed after the final session. First, we present the quantitative data, and this is followed by an analysis of their comments. Their comments showed their thinking process and indicated that they had an awareness of linguistic differences between Japanese and English.

Surveys - Quantitative Data

(i) Summary of responses to the three surveys on the weekly feedback discussions (Appendix 3)

- Across the three questionnaires 80% of students said that they had spent time resolving five or more of their own errors during peer discussions.
- 79% of students found the peer discussions very useful and 20% found them useful.
- In 71% of the peer discussions students said they were able to correct all the highlights.
- 11% of students said that they spent time outside class with others solving essay problems and 69% they said they worked on the essays by themselves.
- Based on their experiences with peer feedback 79% of respondents said they preferred to edit their essays with another student.

(ii) Summary of responses to the final questionnaire (Appendix 4)

- All students said they could understand the highlighting system well
- 88% of students preferred to use the highlighting system (Appendix 1) and discussing the highlighted parts with a partner rather than any of the other options given to them in the survey.
- 88% of students stated that they were usually able to correct all their errors when they wrote their second draft.
- In response to the question ‘Did you ever discuss your errors with another student after

the class?’ only 6 % of students wrote that they always did so and 17% of students often did so.

- 62% of the students said that it made a difference if they worked with a friend.
- When writing their second draft students said they used the following resources most: partner’s comments: 38%. dictionary: 29%; online language resources 25, grammar book, 4%.
- Students preferred to edit their essays as follows: by themselves, 17%; with another student, 25%; both 58%.
- 98% of the students said that they found the sessions with their partners useful or very useful

Student Comments on the Surveys

Many of the students gave written comments about the system on the four feedback forms, which were categorised into three themes: (i) benefits of the system, (ii) problems with the system, and (iii) suggestions for improving the system.

(i) Benefits of the system

Students’ comments are summarized below:

1. The colour coding system was a practical and efficient way to review their writing and understand their errors from another person’s perspective.
2. It was better to get advice about their errors from peers than to try to work them out by themselves
 - *‘My partner had a deeper knowledge of English, so I could learn about many other errors from them’*
3. They learned new grammar and appropriate word choice from their partner because they could see different expressions that they had never used before.
4. They gained confidence by showing their English to others, boosting their self-image as English users.
5. It promoted enjoyment in their writing.
6. They took control of their own learning by taking the initiative and asking the teacher if they still didn’t understand after talking to their partner.
 - *I thought about the problems of my essay before class, then I asked a classmate about the ones I couldn’t solve, so I could use time efficiently in class.*
7. They were motivated to do better next time. One student said that the limited class time meant they tried to reduce the number of errors in subsequent essays.
8. To save their partners’ time, they tried harder to avoid mistakes in their first draft.

Students in Group A are studying to become English teachers and their own teaching methods are likely to be influenced by the activities they have successfully used as learners themselves.

Such learner-centred experiences may supplement the theoretical aspects of their education. Advising their peers on language usage may boost their English language teaching confidence, and as one student noted:

- *'This method makes it easy to discover learners' mistakes and what kind of errors they are. Discussing your writing with your partner is useful for both of you. Correcting mistakes is a way to practice using your knowledge.'*

while another said:

- *'I liked this style the most, so I want my students to give writing feedback in the same style if I become a teacher.'*

(ii) Problems with the system from the students' perspective

Problems with the highlighting system

- *Forgetting which colors show which language mistakes*
1. Problems working with their partner:
 - *It was 'inefficient'.*
 - *Some of the better students felt that they gained little in terms of language development and they did not need a partner to find their errors.*
 - *Weaker students sometimes feel intimidated and suffer low esteem: 'In pair activity, I have to do it with a partner who I don't know well and we can't talk easily.'*
 - *They don't need a partner to find their errors*
 2. Difficulty in identifying errors
 - *Some students may not be able to find their partner's mistakes.*
 - *Difficult to work out the error when many words were highlighted at once.*
 - *Couldn't understand the parts highlighted in gray (style problems).*
 3. Need for teacher's input during the sessions
 - *Not sure if they were editing correctly as they couldn't get 'exact knowledge'.*
 4. Need more response from the teacher after the second draft
 - *Sometimes they could not understand even after consulting with their partner so they needed teacher feedback on the rewrite to know the best answer.*

These comments show how much students value or need the teacher to validate their self-corrections and their written work in general. It should be noted that Group A had teacher feedback on all rewrites, whereas Group B had none due to time pressure.

(iii) Student suggestions for improving the system

1. How best to combine solo correction and peer-feedback discussions?
 - *First, do editing by myself, then if there are problems I can't solve, ask my partner. Finally, I want to ask the teacher.*
 - *Better to check with friends before submitting, not just after submitting.*

These two different perspectives show the importance of individual learning styles, so teachers

need to think how best to cater to them.

2. Talking with a variety of classmates

- *Teachers should reshuffle partners so students can get a variety of opinions.*

3. Spending more class time on this activity

- *It is better to decide how many of each others' errors to check to make the most of the limited time.*

4. Language of communication

Both groups were allowed to use Japanese to discuss the errors and most did. They were clearly focused and engaged in their discussions, but this meant that they were not developing English vocabulary relevant to grammar and the essay writing process. However, one student in Group A thought that students should use English to communicate during this activity.

5. Teaching suggestions

- *The teacher could give an example of sentence correction and the class could give an answer together while referring to it.*
- *Improving style – do not only focus on errors. Students also want to know how to write a stylish essay, not just an accurate one. Discuss together whether there is a different expression that could be used to say the same thing in a better way.*

There was also evidence of deeper engagement by the students with their learning compared to motivation seen in previous writing classes. During the class, students would often call on the teacher for help to solve highlighting problems they could not resolve with their partner.

6. Discussion

Research Questions

1. *Did the students collaborate effectively with their peers?*

Students were very engaged during the activity and during the class they often asked the teacher to solve issues they were unable to solve with their partner. In this sense the system succeeded in getting the students to pay attention to their mistakes and learn from them, in a way they had not necessarily been doing before. The rewrites and their comments show that they were thinking seriously about how to improve their writing for their own satisfaction. This suggests intrinsic motivation for the writing process. It may have helped change their perception of English.

2. *Did the students correct the highlights through peer discussions?*

The data indicated that 79% of the highlights were resolved through peer discussions, particularly grammatical highlights as mentioned in Section 4. The reason why they could not correct all the errors could have been due to their knowledge of English, limited time, or prolonged discussion on one language point, so this does not necessarily indicate how active they were during the sessions.

Discourse features were the most difficult to repair, which suggested that students would

benefit from a list of typical stylistic problems to be handed out at the start of the class to be referred to during the peer discussions.

3. Did the students deepen their understanding of English usage through the activity?

From the three case studies mentioned in Section 4 it was clear that the errors the teacher highlighted and that the students worked on and corrected in the second draft were often also mentioned in the students' feedback comments. This suggests that they were thinking about specific grammatical features and vocabulary usage in an informed way during the discussions and that after class they investigated these features by themselves using various reference resources. This pattern was regardless of the English proficiency of individual students: students throughout the ability range seemed to be able to use the system to improve their writing.

4. Did individual students improve over time?

Regarding the number of highlights they received (Tables 1 & 2), it cannot be said with certainty that improvements were made over time. However, regarding the *types* of errors students were not making the same mistakes they had been making on earlier essays as could be determined from an analysis of individual students' essays. Therefore, there was evidence that students had learned grammatical structures they had previously been unable to use appropriately, even though they may have been making different mistakes in their later essays.

5. How did the students evaluate the system?

As seen in the Results section, the students' feedback forms showed very favourable comments about the system.

Questions raised

The above discussion raises a number of questions about how the system might be improved which are addressed below.

1. Should learners use L1 or L2 for feedback? This depends on the purpose of the task. If it is to develop communication skills then L2, but if the primary aim is to develop writing skills, including better understanding of grammar, vocabulary usage, discourse and other style features, then L1 might be more appropriate – also if time is a factor the discussion would probably be more effective in L1.
2. Some students could only correct certain types of errors but not others. This may be related to ambiguity or inconsistency in the teacher's use of the highlights –when more than one or two words are highlighted (up to clause level) then the student has less information to go on and may re-engineer the clause blindly or randomly. (cf Ferris 1999 on treatable and untreatable errors). Bitchener et al. (2005) explained that their ESL learners were able to significantly improve their accurate use of past tense and definite article because these features were “determined by sets of rules”. The authors

went on to add that the opposite was true for prepositions, which learners could not improve over time because this target was more idiosyncratic.

3. How best to follow up the peer discussions? Individual learners could put up a problem sentence for the rest of the class to consider in pairs and the teacher could give final feedback to the class. Alternatively, the teacher could get learners to show their corrections in subsequent classes or be available for advice outside class if learners are unable to solve particular errors with their partners or by themselves.

Improving the system

- Students need to be given strategies to deal with challenging sentences (avoidance or strategies, compensatory strategies, reviewing & rereading (Schmitt (2010:165-8)
- More time may be needed in class for teaching particular reoccurring grammar points after checking the second drafts.
- The average time taken for the teacher to highlight all student essays each week was about 3 hours, which makes this process very labour intensive. To make the system more viable for busy teachers there is a need to make the process less onerous by streamlining the highlighting process, perhaps by utilising grammar and vocabulary functions in Word or Google docs.

Further study will continue to trial the system in forthcoming classes, including gathering more error data from essays and questionnaires. We also plan to develop reference materials (online or on paper) or to make more systematic use of existing resources (e.g. SCoRE) to better guide students during the peer discussion activities. This could also be complemented by developing a better way of measuring more accurately students' speed of writing over time in order to compare individual learners' performance.

Conclusion

Highlighting puts students' errors in stark relief and in many cases this is enough for them to realise what needs to be repaired, which indicates a gap between declarative knowledge of a rule and the ability to use his knowledge correctly in their writing (procedural knowledge).

The aim of this system was not to be completist but to encourage students to think deeply about their own language and to give them an opportunity to contribute to the learning of their peers. There will always be students who prefer to work by themselves, but the consistent positive results shown in the questionnaires show that collaborative editing is popular.

The system also enhanced teacher-student communication because students were asking more questions to the teacher about their writing than in the past despite (in the case of Group B) not having any instrumental benefit for writing the second draft. There were hints of intrinsic motivation at work but this will be pursued further in the next stage of this research.

Furthermore, the problem raised about potential disparity in English proficiency between pairs of students suggests a need for the teacher to stress how this can be a teaching activity as much as a learning activity. As can be seen from the analysis of the first and second drafts of the three case studies of student essays (Section 4) even the best students were not always able to correct all the highlights, but the process did cause them to consider what kind of problems they were having, gain a deeper understanding of how English works and attempt to repair them in an informed way. The responsibility for developing their language shifts over to the learners themselves and away from the teacher, whose role is to guide them towards independent error correction rather than to act as an arbiter of their accuracy. In other words, this is a wholly communicative and learner-centred approach.

References

- Barker, D. (2010). *An A-Z of Common English Errors for Japanese Learners*. Tokyo: Back to Basics Press.
- Bitchener, J. (2019). The intersection between SLA and feedback research. In K. Hyland & F. Hyland (Eds.), *Feedback in Second Language Writing: Contexts and Issues* (pp85-103). Cambridge: Cambridge University Press.
- Bonilla López, M., Van Steendam, E., Speelman, D. & Buyse, K. (2021). Comprehensive corrective feedback in foreign language writing: the response of individual error categories. *Journal of Writing Research*, 13(1) 31-70. <https://doi.org/10.17239/jowr-2021.13.01.02>
- Carless, D. (2020). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. *Active Learning in Higher Education*. <https://doi.org/10.1177/1469787420945845>
- Chujo, K., Oghigian, K. and Akasegawa, S. (2015). A Corpus and Grammatical Browsing System for Remedial EFL Learners. In A. Leńko-Szymańska & A. Boulton (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning* (pp109-128). John Benjamins Publishing Company: Amsterdam.
- Colpitts, B. D. F. and Howard, L. S. (2018). A comparison of focused and unfocused corrective feedback in Japanese EFL writing classes. *Lingua Posnaniensis*, 60:1, 7-16. DOI:10.2478/linpo-2018-0001
- Elwood, J. A. and Bode, J. (2014). Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System*, 42 333-343. <https://doi.org/10.1016/j.system.2013.12.023>
- Erkan Er, Yannis Dimitriadis & Dragan Gašević (2020): A collaborative learning approach to dialogic peer feedback: a theoretical framework, *Assessment & Evaluation in Higher Education*, DOI: 10.1080/02602938.2020.1786497
- Erkan Er, Yannis Dimitriadis & Dragan Gašević (2021) Collaborative peer feedback and learning analytics: theory-oriented design for supporting class-wide interventions, *Assessment & Evaluation in Higher Education*, 46:2, 169-190, DOI: 10.1080/02602938.2020.1764490
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1-11. [https://doi.org/10.1016/s1060--3743\(99\)80110-6](https://doi.org/10.1016/s1060--3743(99)80110-6)
- Gass, S. (1997). *Input, Interaction, and the Second Language Learner*. Mahwah: Lawrence Erlbaum Associates.
- SCoRE (Sentence Corpus of Remedial English) <https://www.score-corpus.org/en/>
- Sellick, A. and Bury, J. (2018) Japanese University Students' Reflections on Peer Marking in *Shumei University Journal* (秀明大学紀要)

Storch, N. (2019). Collaborative writing as Peer Feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in Second Language Writing: Contexts and Issues* (pp143-161). Cambridge: Cambridge University Press.

Swain, M. (2006). Language, agency and collaboration in advanced second language proficiency. In H. Byrnes (ed) *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. London: Continuum.

Wakabayashi, R. (2008). The effect of peer feedback on EFL writing: focusing on Japanese university students. In *OnCUE Journal* 2:2, 92-110.

Appendix 1: Essay Topics

Essay	Topic	Genre
1. (A=1)	Introduce the person you talked to today in class. You can use notes that you made during the class, and if necessary you can also contact your partner by LINE to check on any information you need.	Reporting
2. (A=6)	Write an essay explaining your experiences learning English.	Expository Opinion
3. (A=3)	Write a description of a place you know well. • Where is it? Why have you chosen this place? Describe the layout. Why do you like it?	Descriptive
4. (A=2)	Choose a famous living person who has achieved something important in his/her life and write a description of him/her. BUT, you must not write his or her name - the person who reads your essay next week will have to guess who the famous person is	Descriptive Opinion Research
5. (A=5)	You've been invited to write a short article about a typical Japanese dish for an international cookery magazine. Think of a dish you know well from one area of Japan and introduce it to an international audience.	Expository Research
6. (A=4)	Do you agree or disagree with the following statement? 'In the world post-Covid 19 our university should continue with a system of online classes.' Explain why you agree or disagree, considering the advantages and disadvantages of face to face and online classes.	Persuasive

Appendix 2 - Highlighting System

Weekly Homework Essays - Editing System

Mistakes

YELLOW = grammar mistake
e. g. He **go** to the supermarket.

GREEN = vocabulary mistake
e. g. I **tripped** to London

BLUE = spelling mistake
e. g. **Their** sitting in the **kichin**.

RED = punctuation mistake
e. g. **w**hat time did you get up **!**
I like football. **B**ecause it is exciting.

GREY = style

There is no English language mistake, but your sentence isn't appropriate.

- The phrase or sentence is too **formal** or too **casual** ('And' ⇔ 'In addition')
- Using **passive voice** (受動態) instead of **active voice** (能動態) or vice-versa
- Using **reported speech** (間接話法) instead of **direct speech** (直接話法) or viceversa.

- Put the most important information at the end of the sentence ('**End focus**').

e.g. (i) I was attacked by a dog on Sunday .

=> (ii) On Sunday I was attacked by a dog .

(In (i) the time is more important than the action. In (ii) the action is more important than the time.)

Other symbols

red font = I have corrected your sentence

I ~~go~~ went to the shops.

[] = word(s) missing

e. g. [] Japanese are very polite. Aetf5e

house = word should be deleted

e. g. I ~~have~~ walked to school yesterday.

Appendix 3: Surveys on Peer Discussions

Essay		Group A			Group B			Mean
		1	6	5	1	2	4	
How many of your own errors did you work on during the peer discussions?	1	0	0	0	3	3	3	1.5
	2	0	0	9	7	7	10	5.5
	3		9	5	3	0	20	7
	4	9	9	9	3	7	0	6.2
	5+	86	81	76	83	83	67	79.3
How useful was the peer discussion?	Very useful	86		81	83	73	70	78.6
	A little useful	14	9	28	13	27	30	20.2
	Not so useful	0	0	0	0	0	0	0
	Not useful at all	0	0	0	3	0	0	0.5
Were there any highlighted errors you couldn't solve during the peer discussions?	YES	-	14	28	43	23	27	27
	NO	-	86	71	57	77	73	72.8
Did you spend time solving essay problems after class? (Peer discussions average time: Group A = 40 mins; Group B = 20 minutes)	Yes, by myself	43	62	57	77	87	87	68.8
	Yes, with another student	0	19	14	17	7	10	11.2
	No	57	19	28	7	7	3	20.2
Based on your experiences with peer feedback, how do you prefer to edit your essays?	By myself	9	19	19	27	23	30	21.2
	With another student	90	81	81	73	77	70	78.7

Appendix 4 Final Survey

	Group A	Group B	All
1 Did you understand the highlighting system well?			
YES	100	100	100
NO	0	0	0
2 Which feedback system do you prefer?			
Highlighting system using different colours (discussing with partner)	90	87	88
Highlighting system using 1 colour (discussing with partner)	0	3	2
Highlighting system using different colours (self-correction)	0	6	4
Highlighting system using 1 colour (self-correction)	0	0	0
Teacher correction of errors	10	3	6
3 Were you usually able to correct your errors when you did your second draft?			
YES	86	90	88
NO	14	10	12
4 Did you ever discuss your errors with another student after the class?			
Always	5	6	6
Often	24	13	17
Occasionally	48	32	38
Never	24	48	38
5 Did it make any difference if you worked with someone who you were already friends with?			
YES	67	58	62
NO	33	42	38
6 When you wrote your second draft, which of the following resources did you use most?*			
Partner's comments	52	29	38
Dictionary	9	42	29
Grammar Book	5	3	4
Online language resources	33	19	25
7. Based on your experiences with peer feedback, how do you prefer to edit your essays?			
By myself	9	23	17
With another student	19	29	25
Both are good	71	48	58
No preference	0	0	0
8 Overall, how useful for your English were the essay feedback sessions with your partners?			
Very useful	81	74	77
A little Useful	14	26	21
Not so useful	5	0	2
Not at all useful	0	0	0