The Effect of Three-Way Corrective Feedback in EFL Writing in a Japanese Junior College Setting

日本の短期大学授業における英文ライティングで三種類の訂正フィードバックを提供することによる効果

Kaoru Mita
Professor, Department of English Communication

Yoshie Kubota
Full-time Lecturer, Department of English Communication

Tomoko Kurita
Part-time Lecturer, Department of English Communication

Yuko Maurer
Part-time Lecturer, Department of English Communication

David Baldwin
Part-time Lecturer, Department of English Communication

Lorna Velia De Vera
Part-time Lecturer, Department of English Communication

Robert Christopher Lavey
Part-time Lecturer, Department of English Communication

三 田　　　薫
英語コミュニケーション学科教授

久保田　佳 枝
英語コミュニケーション学科専任講師

栗 田 智 子
英語コミュニケーション学科兼任講師

マウラー　裕子
英語コミュニケーション学科兼任講師

デイヴィッド・バールドウィン
英語コミュニケーション学科兼任講師

ローナ・ヴェリア・デヴェラ
英語コミュニケーション学科兼任講師

ロバート・クリストファー・レヴィー
英語コミュニケーション学科兼任講師

Summary:

This paper explores how three-way corrective feedback (feedback by automated writing evaluation software, native teacher's feedback and outsourced feedback) affects students' writing for a global student forum, which was conducted for a mandatory general English class at a junior college in Japan. Groups of students wrote four essays and posted on an international site during the term. The results of student questionnaire and post-writing test showed that high proficiency students used immediate indirect feedback more efficiently than low proficiency students; just providing three-way feedback without giving instructions of essay writing did not increase students' fluency in writing; non-native teacher's writing instruction may have had a significant impact on students' fluency.

和文抄録：

　日本の短期大学の必修一般英語コースで、3種類の訂正フィードバックを提供しながら学生グループにエッセイを書かせ、海外サイトに投稿させる活動を 4 回行い、学生の英作文への影響を調べた。その結果、英語力の高い学生は即時・間接フィードバックを有効に活用すること、単に 3 種類のフィードバックを提供するだけでライティング指導を行わない場合には効果が見られないこと、ノンネイティブ教員の指導が事後ライティングテストの語彙数の増加に大きく影響することが分かった。

Key words： computer assisted language learning (CALL), direct/indirect corrective feedback, immediate/delayed corrective feedback, English as a Foreign Language (EFL), second language writing, junior college in Japan, native teacher (NT), non-native teacher (NNT), automated writing evaluation (AWE) software, International Education and Resource Network (*iEARN*)

キーワード： コンピューター支援言語学習、直接／間接訂正フィードバック、即時／遅延訂正フィードバック、外国語としての英語、第二言語作文、日本の短期大学、ネイティブ教員、非ネイティブ教員、機械作文評価ソフトウェア、アイアーン

1. Introduction

　Computer development and technical advancements are about to change the scene of second language writing in college. Error correction, which has long been the heavy workload of instructors, will be supplemented by error checkers, automated writing evaluation (AWE) software, or outsourced corrective feedback provided via the Internet for free or with decreased costs, and their improvements in the efficiency and refinements in computers and software are eye-opening.

Computer and Internet technology has already been used in corporate society on a large scale, and most of business English writing conducted in Japanese companies has been email exchange (小池, 2010), which requires writing correct, prompt, unambiguous and hopefully persuasive messages appropriate for the era of globalization.

That being said, the English writing proficiency of Japanese students is dismal: the Ministry of Education, Culture, Sports, Science and Technology disclosed in a 2015 report that 82.1% of public high school seniors were at the very basic level (CEFR Level A1; STEP Grade 3 to 5) in writing English (文部科学省 , 2015). It may be no exaggeration to say that there is a huge gap between the level that companies generally require and the level that is reflected in students' writing, a problem which should be shared among college instructors to move forward towards solving the issue.

The primary goal of this paper is to explore how three-way corrective feedback affects students' writing for a global student forum. The three-way corrective feedback in the research refers to use of AWE software, native teachers and an outsourced feedback service to groups of first-year junior college students in a mandatory English course; the students' sentences were posted and returned via the Internet all through the three feedback steps. Finally, the complete essays were posted on a global student forum site.

Corrective feedback is one of the most widely researched topics in the second language learning discipline. Also, a growing number of papers have been published on the usefulness of computer-assisted language feedback. What is unique about this paper is that it conducted research on *multiple* uses of corrective feedback which was given to four-proficiency-level college classes aiming to post essays on an international site.

2. Human Feedback vs. Computer-Assisted Feedback

2.1. Students want feedback

Since Truscott (1996), there has been a long debate on whether written error corrections by teachers are effective, but it has been accepted in the literature that the students want feedback. This classic work established it unequivocally in both ESL context (Ferris & Roberts, 2001; Leki, 1991; Radecki & Swales, 1988) and EFL context (Lee, 2004).

In their investigation of student perceptions and preferences concerning teacher feedback in university EFL writing classes in Japan, Elwood and Bode (2014) found that this desire for corrective feedback among Japanese university students seems to be reflected with their positive sentiments and reactions toward feedback on their writing, i.e., Japanese students were "generally very positive" about feedback, and they were "happy" to have received feedback, which "increased their motivation" (p. 339). The same response is observed in other studies on feedback, which showed that students' assessment of teachers' feedback was

"favorable" (Hedgcock & Lefkowitz, 1994, p.156) and "overwhelmingly positive" (Ferris, 1995, p.47). These positive and affective reactions suggest the potential benefits of students from corrective feedback. The studies affirm that with the strong preference for direct, detailed feedback, not only do students want or expect feedback but that they also pay close attention to teachers' feedback. Additionally, students responded that they appreciated the effort and attention given by the teachers and value the feedback on their written errors as it helped them "avoid future mistakes, improve their grammar, and clarify their ideas" (Ferris, 1995, p.48).

2.2. Feedback makes better writers

The usefulness of feedback is well established and supported by many studies. Bitchener (2008) found that the accuracy of students who received written corrective feedback in the immediate post-test outperformed those in the control group and that this level of performance was retained two months later. Most importantly, the study found that written corrective feedback had a significant effect on improving accuracy in the use of two functional uses of the English article system (the use of "a" for the first mention and "the" for subsequent mentions) and that this level of accuracy was retained even after two months (Bitchener, 2008; Ferris & Roberts, 2001).

Results from other studies by Ferris and Roberts (2001) and Jones, Myhill, and Bailey. (2013) which examined feedback on students' linguistic errors and looked at how the teaching of grammar affects students' writing also showed that there was an overall positive impact on the writing of students who received feedback on grammar and it was beneficial for students who were able writers.

2.3. Direct vs. Indirect Feedback

Corrective feedback in writing may be provided either as direct or indirect feedback. Different researchers of L2 writing give similar definitions and explanations for direct and indirect feedback. Direct feedback is given when the teacher specifically provides the correct form (Elwood & Bode, 2014; Ferris & Roberts, 2001) and could include "the crossing out of an unnecessary word/phrase/morpheme, the insertion of a missing word/phrase/morpheme, or the provision of the correct form or structure" (Bitchener, 2008, p.105). Bitchener also adds that direct feedback can come in the form of written meta-linguistic explanations (wherein teachers supply grammar rules and examples at the end of a composition instead of where the error has occurred) and oral meta-linguistic explanation (a mini-lesson where the teacher gives the grammar rules and examples, or one-on-one individual conferences between teacher and student, or conferences between teacher and small groups of students). Meanwhile, indirect feedback is given when the teacher only suggests that an error exists without explicitly

correcting the error (Elwood & Bode, 2014) and leaves the student to solve and correct the error (Bitchener, 2008; Ferris & Roberts, 2001).

Elwood and Bode (2014) raise the merits of giving direct feedback while citing the work of researchers such as Ferris and Roberts (2001), and Timson, Grow, and Matsuoka (1999), which found that teachers and students prefer direct feedback. They point out that direct feedback also reduces the occurrence of misunderstanding, or "confusion about the location or type of the error or about the meaning of error codes used" (Elwood & Bode, 2014, p.334) as this is a problem among lower-proficiency learners (Ferris & Roberts, 2001). However, direct feedback can be challenging as writing detailed comments could be a time-consuming task for the teacher (Chandler, 2003; Lavolette, Polio, & Kahng, 2015).

Providing indirect feedback also has its advantages as described by Lalande (1982) and James (1998) since learners are able to involve themselves in the processes of guided learning and problem-solving which allow them to reflect on their learning that could lead to long-term acquisition as well as the potential to push learners to engage in hypothesis testing, a process that can induce deeper internal processing and promote the internalization of correct forms and structures (Higgins, Hartley, & Skelton, 2002; Vickerman, 2009). Error correction researchers have also pointed out that indirect feedback is more helpful or at least as equally as direct feedback for students to make progress in accuracy over time (Frantzen, 1995; Lalande, 1982; Robb, Ross, & Shortreed, 1986).

2.4. Computer-assisted feedback as a solution to teacher's time constraints

The effectiveness of written corrective feedback is well established as an instrumental part of the advising process. However, due to time constraints, EFL teachers in universities and colleges are generally limited in the amount of grammar correction they can give which would undoubtedly have an impact on the quality of student writing. This is not to mention that it is also becomes a burden for teachers to differentiate their feedback accordingly especially in cases of large-size classes and classes with students who have differing proficiency levels.

One solution to this problem would be to resort to computerized or automated evaluation tools and automated feedback systems, or e-feedback systems offering advice and hints for writing enhancement. These include online grammar and spelling checkers; assessment tools like automated writing evaluation (AWE) programs, also known as Automated Essay Scoring (AES) programs; and writing platforms that offer customizable services. These are vital options that resolve the issues of time constraints and extra burden on teachers and provide additional support for students.

While acknowledging that e-feedback resources are effective tools for evaluation in writing, it must also be noted that computer-assisted feedback are meant to complement written

feedback from the instructor, "a positive addition to, not replacement for, teacher feedback" (Lavolette, et al., 2015, p. 50).

Even as interaction between teachers and students ideally make for effective instruction and lead to improvement in writing, feedback is a "complicated, challenging, and time-demanding activity" and would require effort from instructors and could be difficult to use continuously and consistently (Choi, 2010, p. 5). Computerized or automated feedback systems allow for providing immediate feedback in writing as well as help in lessening the teacher's workload if used properly and with sufficient student training (Lavolette et al, 2015). It was also revealed that over time, there was a high responsiveness from students toward correction that made use of a computer-assisted language learning system.

Choi (2010) wrote on the effectiveness of feedback using three different types: automated or AES feedback, teacher feedback, and a combination of both. He compared the impact of these different types on the quality of students' writing and found that while automated feedback was effective for improving the quality of writing, the combined feedback of both teacher and automated feedback was the most effective among the three kinds of feedback. This is similarly supported by the findings of Chen and Cheng (2008), which show that, despite a lukewarm reception of automated feedback by students, it was perceived "comparatively more favorably" when it was used to facilitate the initial part of the writing process in drafting and revising and then followed by teacher feedback in the latter part of the process. They found that automated feedback with "minimal human facilitation" (p.94) only caused problems for students and made them frustrated about their writing.

By using automated feedback integrated with subsequent feedback from the instructor, Chen and Cheng (2008) assert that while automated feedback presents an opportunity for students to learn to work independently and helps build their confidence in writing, the interaction with the teacher that follows provides the important element of human feedback. Automated feedback may initially provide suggestions for accuracy and organization, but instructors are better able to address students' specific questions and to respond more personally with comments and can thus result in more effective and purposeful writing that has both meaning and form.

In summary, in identifying the value and the merits of combining human and machine capabilities for corrective feedback, it is apparent that using more than one form of feedback might be a key in overcoming teaching limitations for writing and could lead to better student writing output. Corrective feedback is in itself a complicated process and would require patience on top of time and effort from instructors. There are a lot of things to consider when attempting combined automated and human feedback such as: how to use direct and indirect feedback; deciding what type(s) of e-feedback system to use; providing proper training for

students on how to use them; the quantity and manner of teacher feedback; as well as the roles of native and nonnative teachers in giving feedback in an EFL context.

3. Research Question

How does three-way corrective feedback affect students' writing for a global student forum?

4. Method

Participants

The participants were all first-year Japanese female students, mostly aged 18, who took a general English course during the spring term, 2017. They were students from the Department of English Communication. All the students in the junior college were required to take a general English course, namely *Integrated English*, consisting of two classes a week, combined into one course. One class is taught by a non-native (Japanese) teacher, while the other is taught by a native speaker. Students are evaluated on the combined results of both classes. Students took the standardized placement test ELPA in early April 2017, and were placed into four levels from Class A, the highest, to Class D, the lowest. Table 1 shows each class's number of students and the pairs of native teacher (NT) and non-native teacher (NNT).

Table 1. Four Classes of *Integrated English* and the pairs of NT and NNT

Class	Registered number of first grade students	The teacher of Tuesday's Class	The teacher of Friday's Class
Class A	30	NT(A)	NNT(A)
Class B	34	NNT(B)	NT(B)
Class C	32	NT(C&D)	NNT(C)
Class D	28	NNT(D)	NT(C&D)
Total	124		

Note. A, B, C and D in the parentheses mean Class A,B,C,D respectively. NT(C&D) means that the NT taught both Class C and D.

The course is 15-weeks long, 90 min. each. The NT teaches 15 lessons, and the NNT teaches the same number of lessons. In a week, there are two classes of *Integrated English*, one class taught by a NT, and the other taught by a NNT, either on Tuesday or Friday.

The NNT class is focused on increasing grammar knowledge which are expected to lead to the improvement of four skills. Since 2015, the NNT class has used a textbook titled *Basic Grammar in Use* (Cambridge University Press, the English version in 2015, and Japanese version in 2016 and 2017). In the NT class, on the other hand, the focus is on fostering the four skills by collaborating with NT's class.

The Context and Materials

In the Spring term of 2017, *iEARN* (International Education and Resource Network) was introduced as a part of the curriculum of the course. The *iEARN* is "an international education network linking schools of the world online by using Information and Communication Technology (ICT), and by exchanging information and views on the electronic forum on the Web" (Naya, 2006).

iEARN provides a forum page for each of the global projects on the Web. On the forum, students in different countries can post their information on their culture or views on global issues. Posting on the forum of *iEARN* in the English course was expected to provide the students with a good opportunity to raise their awareness of the world and motivate them to write in English to an international audience in a real context. Through several teacher meetings held for discussing how to incorporate *iEARN* into the curriculum, the following four projects and related topics were chosen for essay writing.

1. Holidays (*Holiday Card Exchange*)
2. Food (*Global Food Show & Tell*)
3. Education (*Girl Rising*)
4. Environment (*International Book Club*)

The students were divided into 6 groups in each class, and each group (with 4 to 6 members) was required to join the projects, conduct research and write about the four topics during the Spring term of four-months long (from the middle of April to the end of July).

When the decision was made to incorporate *iEARN* in the curriculum of *Integrated English*, the teachers realized that basic grammatical errors of students' essays should at least be corrected to the comprehensible level before posting on the forums. To overcome this issue while trying not to increase NTs' workload of time-consuming error correction, the teachers tried several systems and services, and finally decided to employ the following:

1. e-feedback by an automated writing evaluation (AWE) software
2. outsourced human corrective feedback

They were expected to replace part of the NT's error correction workload and reduce worries about students' essays before finally posting on *iEARN*'s forum page.

AWE software

The teachers used an AWE software for e-feedback in their classes. It is the AWE software which is designed to provide both instant computer-generated scores and diagnostic feedback

for a submitted essay by a language learner. It was provided free of charge for personal use during the period of current research in 2017. There is a similar system named *Criterion*® by *Educational Testing Service*, which is the most studied intelligent computer-assisted language learning (ICALL) system. The teachers compared both systems, and chose the first mentioned AWE software, mainly for the reason that it was offered free of charge to the public. As *Criterion*, the AWE software for the research provides indirect feedback by locating an error, and gives metalinguistic suggestions of the correct form. Thus, the AWE software feedback is "immediate, indirect, automated, e-feedback" (Lavolette, et al., 2015).

Outsourced Feedback

Outsourced feedback was employed as additional corrective feedback in *Integrated English*. An online English correction service company located in Japan was used, which provided human correction with a charge of around US 10 dollars per 1200 letters (approximately 200 words). The service was for supporting English learners' writing based on the learners' original English sentences; the correction is not aimed at offering a perfect piece of translation, but for helping English learners' writing.

Direct feedback was offered, and it also offered the instructor's comments about grammatical points of the errors. The students could choose either Japanese instructors or native-speaking instructors for correction before they submitted their sentences. Japanese instructors wrote comments in Japanese, and native-speaking instructors in English. Most of the students preferred Japanese instructors because it was easier to understand the instructor's comments written in Japanese.

Each group was given a 200-word correction opportunity per topic of *iEARN*, 800-word correction in total for four topics. If groups' drafts exceeded the 200-word limit, they selected the part they wanted to be checked and submitted it. The corrected sentences and instructor's comments were returned to them online after a couple of days on average. Thus, the outsourced feedback is "delayed, direct, human, e-feedback."

NT's Feedback

Three NTs taught in four classes, in which the NT(C&D) taught two classes C and D in the same way. Their feedback was basically conducted through teacher-student conferences. Table 2 shows the NTs' feedback given in all the four classes, in which *manaba* is the school online platform used for submitting students reports and exchanging messages. As described in Table 2, the NT's feedback for the research is "immediate/delayed, direct/indirect, human, e-feedback/handwritten feedback."

Table 2. NT's Feedback given in Class A, B, C & D

	NT's feedback
Class A	The NT administered feedback through a modified form of teacher-student conferences. The six groups were first asked to post their drafts on *manaba* before their scheduled conference which allowed the NT to make a preliminary check of their work. During the teacher-student conference, the NT met with the members of each group to go over their writing together with the students simultaneously giving comments and suggesting revisions and answering students' questions. These comments and revisions were either handwritten on a printout of the submitted draft or typed directly into *manaba* as a reply to the group's post. Another form of feedback was done by having students view samples of writing from the submitted drafts that were projected on a screen while the NT presented grammatical, punctuation, and spelling errors and gave corrections and advice on how students can improve their drafts. The writing samples were taken from different drafts and chosen based on what the NT observed to be the most commonly committed errors. These were displayed on the OHP for the whole class' perusal but without giving out the names of the group or the writer. During these sessions, the NT encouraged students to take down their own notes especially if they thought that they committed the same errors in their writing.
Class B	Students wrote and re-wrote their presentations on two different online platforms before submitting to *manaba*. Once they were submitted, they were corrected or approved, as some submissions did not need correcting. Direct corrections were given and students were expected to change their versions accordingly. They then had to write up their paragraphs in their notebook and read it to a partner. After the partner read, students came to the front for more error correction and pronunciation practice. Then students were expected to fill in a note card by hand with the finished product. Each group finished at different times so there was never a long line to wait to speak to the teacher. After completing all of these, students stood up, found a place around the room and read facing the wall, reading as slowly and clearly as possible. Some students elected not to participate in the presentations and those students were still expected to practice reading out loud. The NT circled the room offering advice or praise along the way. This would take place in two different class periods during the process. Students then gave poster presentations for a grade. After the presentations, the NT would give notes and offer more pronunciation advice and practice to fill out the 90- minute class time.
Class C & D (the same NT)	Due to time constraints, the writing, review and presentations had to be completed in three classes from start to finish. In the first class, the NT had each team brainstorm, choose a topic from the main subject and divide up the topic individually to put together in their presentations. The NT told them to write in their native language first and then translate to English. This helped the students put their thoughts down in their own words instead of copying information (plagiarizing) from books or the Internet. The NT then walked around from group to group to give advice and keep the students focused on the task. The students were also asked to think about and make individual posters that corresponded to their part of the presentation. They were given A3 size paper to take home and told to make a rough sketch with pencil for homework. Also on their own time or in the NNT class, the students then used the AWE software to check their work and posted it to *manaba*. Before the second class the NT checked *manaba*, printed each group's partial work

and made handwritten notes and corrections. At the beginning of the second class the students were told to continue their writing or sketch making. The NT walked around checking the sketches of the poster homework. If it corresponded to their writing, it was approved, and they were told to use markers to finish. Of course, many of the students drew sketches that had nothing to do with their part or didn't do it at all so the NT told the individuals to revise and continue making their sketches. The teacher then called each group up to the front for a conference to review their writing on the big TV/screen so that each student could see the other group member's parts. The teacher then gave direct feedback individually. The students were then supposed to post to *manaba* again for homework. For the remainder of the second class the students continued making their individual posters for their presentations, practiced reading their part of the presentation and chose their individual presentation order. The NT walked around and gave individual direct feedback on pronunciation and rhythm. For homework, the students were encouraged to memorize their part and "tell" their story instead of reading. They were also told to finish their posters.

Finally, in the 3rd class the students were given time to touch up their posters, practice their presentation parts and ask the NT final questions. The students then gave their presentations. In order to keep the other students engaged, they were given simple evaluation sheets and asked to give each team member an A, B or C grade on voice level, memorization and content. Then they had to give a short-written statement concerning the group's presentation. The NT had each group give themselves a grade and comment on their own evaluation sheets. At the beginning of the next class, the NT gave feedback on the presentations to the whole class and introduced the next topic. While the groups discussed the new topic, the NT called the students up individually to tell them their good points and where they need improvement. Then the whole process started again.

The three ways of feedback introduced above are summarized in Table 3.

Table 3. Types of Feedback Administered in *Integrated English*

Types of feedback	Immediate/ Delayed	Direct/ Indirect	Automated/ Human	e-feedback/ Handwritten
AWE software	Immediate	Indirect	Automated	e-feedback
NT's feedback	Immediate/Delayed	Diect/Indirect	Human	e-feedback/Handwritten
Outsourced feedback	Delayed	Direct	Human	e-feedback

Order of Feedback

Student groups wrote their essays according to the following process, received three ways of feedback, and posted on the *iEARN* website.

1. Research on *iEARN*'s topic (holidays, food, education, environment)
2. Write the first draft and get e-feedback by the AWE software
3. Write the second draft and NT's feedback (either e-feedback, handwritten feedback or teacher-student conferences)

4. Write the third draft and get the outsourced feedback

5. Posting on *iEARN*'s webpage

About five class times (three NNT classes, and two NT classes) were allocated for starting research about the topic, writing a draft, completing it and finally posting the essay on the *iEARN* website. In the NT class, the final lesson assigned for each topic was used for presentation, which was why fewer classes were used for a topic than the NNT's.

The drafts students made were submitted to *manaba* online platform of the school before each step of being submitted to the AWE software, checked by the NT, or corrected by the outsourced correction service. On the platform, all members of the group could both submit and check other members' drafts.

Due to the national holidays and school events which affected the collaborative schedule of *Integrated English*, sometimes the NT could not give students feedback in time, or the outsourced correction did not send the feedback before the class time for the next *iEARN* topic. These problems were solved by extending the date for posting on *iEARN*.

Student's Manual

The Student Manual named *Handbook of English for International Communication* was made by the NNTs of *Integated English*. It contained 1) the lesson schedule of each of the NT and NNT class with specific dates, 2) evaluation criteria of both NT and NTT classes, 3) how to submit sentences to the school's online submission system *manaba*, 4) how to use the AWE software, 5) how to use the outsourced correction service, 6) how to post essays on *iEARN*'s website, 7) basic grammar for making essays, 8) how to brainstorm and make a mind map, 9)how to make five-structure essays, 10) samples of four essays to be posted on *iEARN*'s website and 11) how to use an online translation system properly.

In the NNT's class, each group was given a "Progress Sheet" every week for recording what they have finished. The group was required to write down in the sheet whether assigned units of the textbook *Basic Grammar in Use* were completed, whether they finished several processes of writing and posting essays to *iEARN*'s website, and what they learned from grammar and writing in the class. The sheet was returned to the NNT during the class.

Teacher's Manual and Teacher Meetings

Both NNTs and NTs had meetings several times before and after the beginning of the new *Integrated English* in April 2017. They discussed how to use the new writing support systems, how to incorporate *iEARN*'s topics into the class, in addition to making "integrated" lesson plans of both NTs and NNTs. Since the NTs planned to use grammatical points taught in the

previous NNT class for their class activities, the NNTs and NTs collaboratively prepared the teacher's manual, in which the detailed contents of each class were written with the specific date.

Although it was ideal to start the course with the NNT's class because it was possible to show students the topic of *iEARN* and had students do basic research about it before passing the class on to the NT's class, this "NNT to NT" order was easily broken because of the timing for the initial class (either Tuesday or Friday), or inserted national holidays and school events. To compensate for this problem, not only the teacher's manual but communication between the pair NT and NNT was important.

The instruction of how to use *manaba*, the AWE software, and the outsourced correction service was conducted by the NNTs. They frequently had meetings and exchanged emails to learn how to use the systems and how to solve unexpected troubles.

Procedures

Pre- and post- grammar tests were administered in the initial and final NNT classes. The tests were 100 multiple-choice questions with a forty minute time limit. The questions were based on the questions in the textbook *Basic Grammar in Use*. All the participants used a computer to take the pre- and post- grammar tests, and the test-scoring system of *manaba* processed and showed the scores immediately after the exam.

Pre- and post- writing tests were administered in the initial and final NT classes. It was a ten-minute writing test with the question: "*Write about your hometown, giving examples about the local people, food, festivals, famous places, etc.*" The test papers were collected during the class, and the word number of each test was counted by students, checked and reported by the NTs using the Excel sheet.

The student questionnaire was administered in the final NNT class. All the participants used a computer to answer the questionnaire consisting of quantitative questions and open-ended questions.

The rubric of *Integrated English* was administered for students' self-assessment in the initial and final NNT classes.

SPSS Version 22.0 for Windows was used for all statistical tests.

5. Results

Scores of pre- and post- grammar tests are shown in Table 4.

Table 4. Means of Pre- and Post- Grammar Tests in each class

	Pre	Post	Difference
Class A (n=26)	63.50	68.88	5.38
Class B (n=31)	53.48	62.61	9.13
Class C (n=30)	47.40	55.13	7.73
Class D (n=26)	41.38	46.27	4.88

Note. The test score is from 0 (Min.) to 100 (Max.).

The result of the t-Test showed that the pre and post scores differed significantly in all four classes (Class A: $t(25)$=4.61, p<.001. Class B: $t(30)$=5.54, p<.001. Class C: $t(29)$=4.83, p<.001. Class D: $t(25)$=3.80, p<.01.). The effect size (d) of each class is in Table 5.

Table 5. Effect sizes & Correlations between Pre- and Post- Grammar Tests

	d	r
Class A (n=26)	0.90	0.92
Class B (n=31)	0.99	0.65
Class C (n=30)	0.85	0.63
Class D (n=26)	0.75	0.81

Note. The effect size is Cohen's d. P-value < 0.01.

Results of pre- and post- writing tests are shown in Table 6.

Table 6. Means of Pre- and Post- Writing Tests (ten-minute writing)

	Pre	Post	Difference	> 100 (Pre)	> 100(Post)
Class A (n=27)	57.26	104.89	47.63	0	13
Class B (n=27)	41.78	64.00	22.22	0	1
Class C (n=29)	39.31	89.21	49.90	0	11
Class D (n=24)	41.50	56.08	14.58	0	1

Note. ">100"= number of students who wrote more than 100 words.

The t-Test showed that the pre and post word numbers differed significantly in all four classes (Class A: $t(26)$=8.21, p<.001. Class B: $t(26)$=6.55, p<.001. Class C: $t(28)$=13.29, p<.001. Class D: $t(24)$=4.30, p<.001.). There were 13 students in Class A, and 11 students in Class C, who wrote more than 100 words in 10 minutes in the Post Writing Test, in contrast to one student each in Class B and D. The effect size (d) of each class is in Table 7.

Table 7. Effect sizes & Correlations between Pre- and Post- Writing Tests
(ten-minute writing)

	d	r
Class A (n=27)	1.58	0.45
Class B (n=27)	1.25	0.55
Class C (n=29)	2.46	0.63
Class D (n=24)	0.87	0.52

Note. The effect size is Cohen's d. P-value <0.05 for Claas A and <0.01 for Class B, C, & D.

There was no correlation between the difference of Pre- & Post- word number of the writing test and the difference of Pre- & Post- score of the grammar test in all four classes.

Table 8 shows students' satisfaction of NNT's Class. From Table 8 to Table 16, questions were designed purposely not to include "Neutral" as a choice option.

Table 8. Students' Satisfaction of NNT's Class

"Was NNT's class good?"	Not good	Not very good	Moderately good	Very good
Class A (n=27)	0	0	6 (22.2%)	21 (77.8%)
Class B (n=31)	1 (3.2%)	1 (3.2%)	18 (58.1%)	11 (35.5%)
Class C (n=29)	0	0	19 (63.3%)	10 (33.3%)
Class D (n=25)	0	0	7 (28.0%)	18 (72.0%)
Total (N=112)	1 (0.9%)	1 (0.9%)	50 (44.6%)	60 (53.6%)

All four NNT classes were perceived as very satisfying. The combined average percentage of "very good" and "moderately good" of the four classes was 98.2%.

The positive responses to the open-ended question about NNT's class included comments on grammar instruction and teamwork: the NNT taught grammar from basics, so they could understand it very well (Class A, B, C, D); whenever they asked what they didn't understand, the NNT explained it very kindly (Class A, B, C, D). Also, they commented that the team activities were worth trying: they enjoyed group activities of discussing and doing research, which strengthened their teamwork; they could learn a lot from the group activities, not only English but various things, which were fun (Class A, B, C, D). There were positive comments on NNT's class as covering what they could not understand in the NT class: they could ask what they couldn't understand in the NT class (Class A, C).

Table 9 shows students' satisfaction of NT's Class.

Table 9. Students' Satisfaction of NT's Class

"*Was NT's class good?*"	Not good	Not very good	Moderately good	Very good
Class A (n=26)	0	2 (7.7%)	19 (73.1%)	5(19.2%)
Class B (n=31)	2 (6.5%)	1 (3.2%)	16 (51.6%)	12(38.7%)
Class C (n=29)	0	0	16 (55.2%)	13(44.8%)
Class D (n=25)	1 (4.0%)	4 (16.0%)	8 (32.0%)	12(48.0%)
Total (N=111)	3 (2.7%)	7 (6.3%)	59 (53.2%)	42(37.8%)

All four NT classes were perceived as very satisfying. The combined average percentage of "very good" and "moderately good" of the four classes was 91.0%.

　The positive responses to the open-ended question about NT's class included comments on their satisfaction by challenging: it was so hard for them to understand what the NT spoke, but it improved their listening skills (Class A, B, C, D); they could learn how to make presentations all in English (Class A, B, C, D); the NT prepared examples for presentation activities, which was helpful (Class A); the NT taught them how to pronounce words while they did their presentation practices, which was very helpful (Class B, C, D).

　Table 10 shows students' perceived usefulness of the sequence of three-way feedback (the AWE software, the NT and outsourced correction).

Table 10. Students' Perceived Usefulness of the Sequence of Three-Way Feedback

"*Was the sequence of your draft checked three ways (The AWE software, your NT and outsourced correction) useful for improving your English?*"	Not useful	Slightly useful	Moderately useful	Greatly useful
Class A (n=27)	0	2(7.4%)	17(63.0%)	8(29.6%)
Class B (n=31)	3(9.7%)	2(6.5%)	22(71.0%)	4(12.9%)
Class C (n=29)	1(3.4%)	4(13.8%)	16(55.2%)	8(27.6%)
Class D (n=24)	1(4.2%)	3(12.5%)	14(58.3%)	6(25.0%)
Total (N=111)	5(4.5%)	11(9.9%)	69(62.2%)	26(23.4%)

The sequence of three-way feedback was highly perceived as useful in all four classes. The combined average percentage of "very useful" and "moderately useful" of the four classes was 85.6%.

　Students' positive comments on perceived usefulness of the sequence of three-way feedback: they could clearly understand what was wrong in their sentences (Class A, B, C, D); by taking several steps of correction, they felt they were able to write with correct grammar (Class A, B, C, D); they felt they were becoming able to write more and more native-like English by going through the steps, which helped for improving their English (Class A, B, C, D). Since it was their first time to experience several steps of corrections, it was difficult to get accustomed to, but worthwhile (Class A, B, C, D).

However, three students commented that it was not useful because they relied on the leader of the group and didn't try by themselves (one student from Class C, and two students from Class D).

Table 11 shows students' perceived usefulness of the AWE software.

Table 11. Students' Perceived Usefulness of the AWE software

"Was the AWE software useful for improving your English?"	Not useful	Slightly useful	Moderately useful	Greatly useful
Class A (n=26)	0	2(7.7%)	15(57.7%)	9(34.6%)
Class B (n=31)	4(12.9%)	2(6.5%)	22(71.0%)	3(9.7%)
Class C (n=28)	3(10.7%)	3(10.7%)	17(60.7%)	5(17.9%)
Class D (n=23)	0	10(43.5%)	8(34.8%)	5(21.7%)
Total (N=108)	7(6.5%)	17(15.7%)	62(57.4%)	22(20.4%)

Class A was the highest in the combined percentage of "very useful" and "moderately useful" on perceived usefulness of the AWE software, and Class D was the lowest: Class A (92.3%), Class B (80.6%), Class C (78.6%), Class D (56.5%). This is the order of English proficiency according to the standardized placement test conducted before the course started.

Students' positive comments on perceived usefulness of the AWE software: it was useful because it gave them immediate correction, and they could try again and again (Class A, B); they were able to understand their weak point in English (Class A, B); it was good to use the AWE software for correcting basic mistakes before checked by the NT (Class C).

On the other hand, some students could not use it well. Students' negative comments on perceived usefulness of the AWE software: they couldn't understand the correction because it was written in English (Class B, C, D); they just copied and pasted their sentences (Class D); there were troubles while using the computer, and they couldn't use the software well (Class A, B, D); they didn't use it because the other members used it (Class D).

Table 12 shows students' perceived usefulness of the NT's correction on their essays.

Table 12. Students' Perceived Usefulness of the NT's Correction

"Was your NT's correction useful for improving your English?"	Not useful	Slightly useful	Moderately useful	Greatly useful
Class A (N=26)	0	2(7.7%)	9(34.6%)	15(57.7%)
Class B (N=31)	1(3.2%)	0	15(48.4%)	15(48.4%)
Class C (N=28)	0	1(3.6%)	15(53.6%)	12(42.9%)
Class D (N=24)	0	6(25.0%)	11(45.8%)	7(29.2%)
Total (N=109)	1(0.9%)	9(8.3%)	50(45.9%)	49(45.0%)

The NT's feedback was highly perceived as useful in Class A, B and C, but considerably low in Class D: Class A (92.3%), Class B (96.8%), Class C (96.5%), Class D (75.0%).

Students' positive comments on perceived usefulness of the NT's correction: they found that some expressions they used were incorrect or strange from the NT's point of view (Class A, B, C, D); they learned how to use words appropriately according to the context (Class A, B, C); they solved all the grammar problems because the NT directly corrected them (Class A, B, C, D); the NT corrected their pronunciation after correcting grammar, which was useful (Class B); they felt the NT's correction is reliable (Class C).

Students' negative comments on perceived usefulness of the NT's correction: it took a long time before they got the correction, so they forgot what were the corrected points (Class A); their sentences haven't been corrected a lot (Class C, D); there was no explanation of why the sentences were wrong (Class B, D).

Table 13 shows students' perceived usefulness of the outsourced correction.

Table 13. Students' Perceived Usefulness of the Outsourced Correction

"*Was the outsourced correction useful for improving your English?*"	Not useful	Slightly useful	Moderately useful	Greatly useful
Class A (n=27)	1(3.7%)	4(14.8%)	14(51.9%)	8(29.6%)
Class B (n=31)	3(9.7%)	3(9.7%)	18(58.1%)	7(22.6%)
Class C (n=29)	1(3.4%)	5(17.2%)	13(44.8%)	10(34.5%)
Class D (n=23)	0	4(17.4%)	12(52.2%)	7(30.4%)
Total (N=110)	5(4.5%)	16(14.5%)	57(51.8%)	32(29.1%)

The combined percentage of "very useful" and "moderately useful" was around 80% in all four classes: Class A (81.5%), Class B (80.6%), Class C (79.3%), Class D (82.6%).

Students' positive comments on perceived usefulness of the outsourced correction: it was useful for improving writing (Class A, B, C, D); they had an opportunity of checked their sentences objectively through outsourced correction (Class A); they liked that their sentences were double checked in addition to the NT's correction (Class A, B); it was useful as the final stage of correction before sent to *iEARN* (Class A, B); they liked the comments the outsourced correction instructors gave them (Class A, B, C, D); they could use the comments for next writing (Class A, B, C, D); thanks to the instructors' comments written in Japanese, it was easy to understand the corrected points (Class A, B, C, D); by comparing pre- and post- sentences, they could learn a lot both about grammar and sentence structure (Class A, B, C, D).

Students' negative comments on perceived usefulness of the outsourced correction: it was basically useful, but the limitation of word number for posting was a bother (Class B); the replies were sometimes delayed (Class A); they were not sure about it because they didn't use

it (Class A, B, D); they think it was not useful for their English because they haven't checked the corrected essays, just copied and posted on *iEARN* (Class A, C); they didn't understand how to use the outsourced correction website (Class A, B, C, D); they thought it wasn't necessary to be double checked because the NT's correction was enough (Class A).

Table 14 shows students' perceived usefulness of *iEARN*.

Table 14. Students' Perceived Usefulness of *iEARN*

"*Was iEARN useful for improving your English?*"	Not useful	Slightly useful	Moderately useful	Greatly useful
Class A (n=27)	2(7.4%)	8(29.6%)	10(37.0%)	7(25.9%)
Class B (n=31)	3(9.7%)	7(22.6%)	17(54.8%)	4(12.9%)
Class C (n=29)	1(3.4%)	5(17.2%)	19(65.5%)	4(13.8%)
Class D (n=24)	0	8(33.3%)	10(41.7%)	6(25.0%)
Total (N=111)	6(5.4%)	28(25.2%)	56(50.5%)	21(18.9%)

The combined average percentage of "very useful" and "moderately useful" of the four classes was 69.4%. Class C was the highest, and Class A was the lowest: Class A (63.0%), Class B (67.7%), Class C (79.3%), Class D (66.7%).

Students' positive comments on perceived usefulness of *iEARN*: it was inspiring to post a comment on the site where people in the world could read it (Class A, B, C, D); they thought it was good since they could read the comments posted by students all over the world, which were useful when they wrote our essays (Class A, B, C, D); they could be very conscious of writing natural English because it was to be read by people in the world (Class A, B, C, D); comments by students from other countries were fun to read (Class B, C); It was tough to write essays by themselves, but they learned a lot of vocabulary, and improved their English through the experience (Class B); since their essay would be read by people outside of Japan, they could become more serious and work very hard with other team members in writing English (Class B); they could learn how other countries' English grammar was actually used (Class C); they could experience various cultures and traditions of other countries (Class C); they could spread their country's culture and tradition to people in the world (Class C); they were happy because they got connected with people abroad (Class C); they thought *iEARN* was useful just because they learned a lot from writing essays to be posted on *iEARN* (Class A); it was a good opportunity for them because they set a goal for posting, and used their teamwork to achieve the goal (Class D); they improved skills of summarizing Japanese sentences into English (Class D); they got accustomed to reading English sentences (Class D).

Students negative comments on perceived usefulness of *iEARN*: It was a pity that they couldn't get a reply (Class A, B, D); after posting essays, they hardly checked the page (Class

A, D); it was difficult to use the website (Class A, B, C, D); they haven't posted essays by themselves, so they didn't know (Class A); only the leader worked for it (Class B); it was not related with improving English (Class A, C, D).

Table 15 shows students' perceived self-confidence in English by taking the course *Integrated English*.

Table 15. Students' Perceived Self-Confidence in English by Taking the Course

"*Did you increase your self-confidence in English by taking this course?*"	Not increased	Slightly increased	Moderately increased	Greatly increased
Class A (n=27)	0	3(11.1%)	21(77.8%)	3(11.1%)
Class B (n=27)	1(3.7%)	3(11.1%)	20(74.1%)	3(11.1%)
Class C (n=28)	0	2(7.1%)	22(78.6%)	4(14.3%)
Class D (n=23)	1(4.3%)	2(8.7%)	13(56.5%)	7(30.4%)
Total (N=105)	2(1.9%)	10(9.5%)	76(72.4%)	17(16.2%)

The combined average percentage of "greatly increased" or "moderately increased" their confidence in English by taking the course was 88.6%, which means highly perceived in all four classes, and Class C was the highest: Class A (88.9%), Class B (85.2%), Class C (92.9%), Class D (87.0%).

Table 16 shows students' self-evaluation of effort level in studying for the course.

Table 16. Students' Self-Evaluation of Effort Level in Studying for the Course

"*Evaluate your effort level in studying for this course!*"	Didn't make efforts	Slightly made efforts	Moderately made efforts	Greatly made efforts
Class A (n=27)	0	1(3.7%)	11(40.7%)	15(55.6%)
Class B (n=27)	0	2(7.4%)	19(70.4%)	6(22.2%)
Class C (n=28)	0	1(3.6%)	22(78.6%)	5(17.9%)
Class D (n=23)	1(4.3%)	4(17.4%)	11(47.8%)	7(30.4%)
Total (N=105)	1(1.0%)	8(7.6%)	63(60.0%)	33(31.4%)

The combined percentages of "greatly made efforts" and "moderately made efforts" in Class A, B and C were very high, but not high in Class D: Class A (96.3%), Class B (92.6%), Class C (96.4%), Class D (78.3%).

Table 17 to 20 below show correlation coefficients between the perceived usefulness of the three types of feedback and *iEARN* for improving students' English, from Class A to Class D in order.

Table 17. Correlation Coefficients between the Perceived Usefulness of the Three Types of Feedback and iEARN for Improving Students' English (Class A)

Class A	AWE software	NT's Feedback	Outsourced Feedback	iEARN
AWE software	-	.440*	.122	.236
NT's Feedback		-	.359	.467*
Outsourced Feedback			-	.501**
iEARN				-

*Note. n=26. **p < .01; *p < .05.*

Table 18. Correlation Coefficients between the Perceived Usefulness of the Three Types of Feedback and iEARN for Improving Students' English (Class B)

Class B	AWE software	NT's Feedback	Outsourced Feedback	iEARN
AWE software	-	.366*	.609**	.702**
NT's Feedback		-	.455*	.468**
Outsourced Feedback			-	.825**
iEARN				-

*Note. n=31. **p < .01; *p < .05.*

Table 19. Correlation Coefficients between the Perceived Usefulness of the Three Types of Feedback and iEARN for Improving Students' English (Class C)

Class C	AWE software	NT's Feedback	Outsourced Feedback	iEARN
AWE software	-	.282	.757**	.865**
NT's Feedback		-	.222	.398*
Outsourced Feedback			-	.799**
iEARN				-

*Note. n=28. **p < .01; *p < .05.*

Table 20. Correlation Coefficients between the Perceived Usefulness of the Three Types of Feedback and iEARN for Improving Students' English (Class D)

Class D	AWE software	NT's Feedback	Outsourced Feedback	iEARN
AWE software	-	.619**	.465*	.328
NT's Feedback		-	.266	.604**
Outsourced Feedback			-	.553**
iEARN				-

*Note. n=23. **p < .01; *p < .05.*

6. Discussion

Our research question was how three-way corrective feedback affected students' writing for a global student forum. The data shows different results according to the difference of proficiency and NNT's instruction. First, the perceived difference of three-way feedback coming from the levels of proficiency will be discussed. Next, the NNT's instruction which caused the different results in the word number of post-writing test will be demonstrated.

The three-way corrective feedback (AWE software, NT's corrective feedback and outsourced corrective feedback) was provided to all four-proficiency-level classes. The three-way feedback was perceived useful in all the classes with more than 85% positive responses (Table 10). However, the perceived usefulness of each corrective feedback was different, which will be unfolded below.

The three-way corrective feedback

High-proficiency (HP) students used immediate indirect feedback by the AWE software more efficiently than low-proficiency (LP) students. Table 11 shows that perceived usefulness of the AWE software was ranked in order from the top (Class A, 92.3%) to the bottom (Class D, 56.5%). The HP students commented it useful because it gave them immediate indirect correction so that they could try revising again and again, or they could learn weak points of their own English. On the other hand, the LP students generally struggled in using it because they could not understand both how to use the software and the indirect corrective feedback provided in English, which caused some students to abandon learning from the feedback and just do copying and pasting.

NT's correction (both immediate & delayed, direct & indirect feedback) was perceived useful in the higher three classes (Class A, B and C) with over 92% positive responses, but not as high in Class D (75.0%) as shown in Table 12. The students of the higher three classes commented that it was useful because they could learn reliable expressions appropriate for the context from the NT's point of view, while some LP students could not understand why their original sentences were wrong after they received direct feedback.

The outsourced corrective feedback was positively perceived by around 80% positive responses (Table 12), with the highest in Class D (82.6%). The HP students perceived that it was an opportunity to be checked their sentences objectively, or double checked in addition to NT's correction before posted on the international student exchange website, while the LP students perceived it useful because the correction was attached with grammar instruction written in Japanese. Some students commented negatively because it was delayed feedback, and the process of uploading and viewing the result was complicated.

Nonetheless, the correlation of the perceived usefulness between outsourced feedback and

AWE software is significant in Class B (.609**), Class C (.757**), and D (.465*), which indicates that there were certain students who used both AWE software and outsourced feedback effectively Table 18, 19, 20).

The NNT's instruction

The results of the pre- and post- exams show that all the four classes increased significantly in both post-grammar and post-writing test compared with the pre- tests (Table 4, 6). The lessons of all four classes were perceived as very satisfying by students with more than 90% positive responses in both NNT and NT classes (Table 8, 9). However, the results of the pre- and post- 10-min. writing tests showed a clear distinction between Class A&C and Class B&D (Table 6). While there was no student who wrote over 100 words in the pre-writing test in all four classes, in the post-writing test, more than 10 students wrote over 100 words in both Class A & C, in contrast to only one student each in Class B & D. The reasons of such a difference will be discussed below.

As shown in Table 2, the NT feedback was conducted well in all four classes, whereas some of the instructions by NNTs were provided in Class A & C, but not in Class B & D. The NNT's instruction in Class A & C had the following similarities:

The similarities of NNT's instruction in Class A and C:
1) Emphasizing the difference of word order between Japanese and English
2) Encouraging students to use their own vocabulary to express ideas (not to rely on automatic translation system)
3) Emphasizing the importance of paragraph structure (a topic sentence and specific details) in writing

Particularly, Class C showed the highest increase in the pre- and post- writing test word number (from average 39 words (Pre) to 89 words (Post)) even though the class level was the second from the bottom. Moreover, Class C showed the highest score (92.6%) in the combined percentages of "greatly made efforts" and "moderately made efforts" in studying for the course (Table 16), and the highest score in the combined average percentage of "greatly increased" or "moderately increased" their confidence in English by taking the course. The instruction conducted in Class C will be examined in detail to find out the reason of the outstanding results.

In Class C, the NNT and the NT had a good collaboration to have the students complete the *iEARN* writing activities. However, it does not mean that both teachers matched exactly what they taught about in advance. The NNT taught how to write English while the NT helped

correct their writing and encouraged students to write more information from the perspective of native English speaker.

In terms of writing approaches for instruction, the NNT of Class C combined genre approach and process approach towards writing. The genre approach is text-focused, which emphasizes the social nature of writing, while the process approach is a writer-centered process, which emphasizes its cognitive aspects. Through combining different approaches, the strengths of one might complement the weakness of the other (Hyland, 2003). Namely, she taught first about the structure using a model text and had the students understand the genre. Then, she had them work on writing focusing on the process including planning, drafting and publishing. Students were encouraged to revise their own writing and publish it as a post on the site of iEARN.

Why did the students in Class C increase the word number of the post-writing test even though they were at the level for remedial English education, one of the low proficiency classes? As for effective instruction, the NNT taught some specific writing strategies when the students had difficulties in L2 writing. *iEARN* writing activities with such topics as education or environment were challenging for the students in Class C because they had a lack of basic knowledge on grammar and vocabulary. They had a huge gap between L1 and L2 writing skills. High proficiency in L1 may not be necessarily an advantage for L2 writing (Hyland, 2003). At the beginning, they often used automated translation system to translate their text written in L1 into English and it turned out the translated text did not make sense. They seemed to feel frustrated by failing to convey what they would like to say in Japanese. The NNT of Class C walked around the groups to check, asked what they would like to convey and helped them to write. She thought it was key for them to write intelligible text using their own English vocabulary. She also taught them about the differences between Japanese and English in writing.

More specifically, there are four factors which seem to be effective in the instruction of Class C. First, the NNT taught English word order SVO, which is different from Japanese word order SOV, as a part of grammar instruction. Second, she asked the students to simplify their ideas in Japanese as if they explained it to children in simple Japanese so that they could express themselves with familiar English vocabulary without using automated translation system. Third, she taught English paragraph structure with a topic sentence, reasons and examples, which is different from Japanese writing structure. Forth, she encouraged them to elaborate on the content by considering of the audience who did not know about the topic and who had different cultural backgrounds. Hyland (2003) mentioned, "Effective L2 writing instruction can make schemata differences explicit to students, encouraging consideration of audience and providing patterns of unfamiliar rhetorical forms (p. 50)."

Also, a factor which was not in the NNT's explicit instruction may have increased their word number in the post-writing test, i.e., *iEARN*, which is the site for collaborative learning with the world. Although the NNT did not teach them about *iEARN* intentionally, she might have influenced the students' perception on writing essays as she knew more about it than other teachers because she had been involved in *iEARN* projects. Therefore, writing may have been perceived not just as a writing assignment, but as a part of the *iEARN* project, for meaningful communicative purposes. That was why they worked hard to post their text at its best on *iEARN* forum, which is open to the audience around the world. In fact, Table 14 shows that students' perceived usefulness of *iEARN* for improving their English was the highest with 79.3% positive responses in Class C. Also, Table 19 shows that the correlations of the usefulness of *iEARN* with that of all the three-way feedback are significant in Class C.

Writing is a purposeful and communicative activity. Therefore, writing tasks should not simply emphasize linguistic ability for formal accuracy, but be situated in meaningful contexts as genuine communicative purpose (Hyland, 2003). It can be said that the students of Class C viewed the *iEARN* writing activity as genuine cross-cultural communication, as expressed in the comment: "*I could experience various cultures and traditions of other countries. I could also spread our country's culture and tradition to people in the world.* (Class C)"

According to Table 18 to 20, correlation coefficients between the perceived usefulness of the three types of feedback and *iEARN* for improving students' English, there is the strongest correlation in Class C between AWE software and *iEARN*. It seems that hardworking students in Class C increased their writing skills through the revising process by utilizing AWE software for posting essays on *iEARN*, which provided them with a meaningful opportunity to improve writing skills. Thus, the several factors such as instruction, feedback and *iEARN* might contribute to increasing the word number in the post-writing test in Class C.

Class B and D did not show as much increase of word number in the post-writing test as Class A and C. There are following similarities in the NNT's instruction of Class B and D.

    <u>The similarities of NNT's instruction in Class B and D:</u>
        1) Focusing on grammar instruction by using a grammar textbook for the course
        2) Not giving special writing instruction in class except for teaching how to use various
           ICT tools

The Class B students increased their scores most among the four classes in the post grammar test (Table 4). However, the improvement in grammar knowledge did not translate into the increase in the number of words in the writing test. The NNT focused on finishing the first 39 units of the textbook which was the term's goal. Focusing more on grammatical

accuracy could have prevented the students from writing more. Another reason could be that there were less exercises to write in Class B compared with Class A and Class C. The NNT did not teach students to focus on how to construct English sentences, i.e. the five sentence structures (SV, SVC, SVO, SVOO and SVOC).

The NNT of Class D also focused on instruction of the grammar covered in the course textbook. The NNT asked the students to post at least one question about quizzes they could not understand in the textbook with comments on the reason why. The posted questions from the students were surprisingly basic that the NNT thought it necessary to teach only the very basic level of grammar. She explained the basic grammatical points in detail, and the students commented that they could understand the NNT's explanation very well. However, Class D showed the lowest increase in both the post-grammar and post-writing test. This coincides with the results of research by 酒井 (2015) on remedial English education for college students that demonstrates that relearning the contents of middle and high schools does not raise substantial effect. Subsequently, the students may have not acquired how to write English sentences, how to understand the meanings of indirect feedback, and in the end, abandoned learning opportunities just by copying and pasting the sentences in each stage of three-way feedback. As a matter of fact, Table 16 shows that students' self-evaluation of effort level in studying for the course was the lowest in Class D (78.3%). The result of Class D indicates that the use of feedback by LP students should be treated cautiously, which is also pointed out by Lavolette et al. (2015) and Cavaleri & Dianati (2016).

There are several limitations in this study because we examined only the word number of students' writing, which indicates fluency. Accuracy and complexity of the writing should be also examined in the future research.

7. Conclusion

This paper explored how three-way corrective feedback affected students' writing for a global student forum. The three-way corrective feedback (feedback by AWE software, NTs and an outsourced feedback service) was provided to groups of first-year junior college students in a mandatory English course.

The results showed that 1) high proficiency students used immediate indirect feedback by the AWE software more efficiently than low proficiency students, 2) just providing three-way feedback without giving instructions of essay writing did not increase students' fluency in writing, 3) the NNT's writing instruction may have had a significant impact on students' fluency even with low proficiency students. In addition, giving an opportunity to post essays for meaningful communicative purposes may have enhanced self-confidence in English. With the development of computer and the Internet technologies, some might expect that there will

be decreased teachers' role in the future. The survey revealed the contrary results, and the role of NNT was highlighted as a key to affect students' writing skills and self-confidence in English.

References

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102-118.

Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, *10*(1), A223-A236.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*, 267-296.

Chen, C. F. E., & Cheng, W. Y. E. C. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, *12*(2), 94-112.

Choi, J. (2010). The impact of automated essay scoring (AES) for improving English language learner's essay writing. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3437732).

Elwood, J. A., & Bode, J. (2014). Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System*, *42*, 333-343.

Ferris, D. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly*, *29*, 33–53.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, *10*(3), 161-184.

Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *Modern Language Journal*, *79*, 329–344.

Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, *3*, 141–163.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, *27*, 53-64.

Hyland, K. (2003). *Second language writing*. New York: Cambridge University Press.

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.

Jones, S., Myhill, D., & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students" writing. *Reading and Writing*, *26*(8), 1241-1263.

Lalande, J. F., II (1982). Reducing composition errors: An experiment. *Modern Language Journal*, *66*, 140-149.

Lavolette, E., Polio, C., & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. Language, *Learning & Technology*, *19*(2), 50-68.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, *13*(4), 285-312.

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals*, *24*, 203–218.

Naya., Y.(2007). "How Intercultural Disaster Reduction Education Change Students: A Case Study of an Evening Course Senior High School in Hyogo, Japan". IWIC'07 Proceedings of the 1st international conference on Intercultural collaboration, p.368-381, Kyoto, Japan. Retrieved from https://pdfs.semanticscholar.org/77e2/43bf083900741101a92b49dddc6f0df57ed2.pdf

Radecki, P., & Swales, J. (1988). ESL student reaction to written comments on their written work. *System*, *16*, 355–365.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, *20*, 83–93.

Timson, S., Grow, A., & Matsuoka, M. (1999). Error correction preferences of second language learners: a Japanese perspective. *JACET Bulletin*, *30*, 135–149.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*, 327–369.

Vickerman, P. (2009). Student perspectives on formative peer assessment: An attempt to deepen learning? *Assessment and Evaluation in Higher Education*, *34*(2), 221-230.

文部科学省 (2015) 平成 27 年度 英語力調査結果（高校 3 年生）の速報（概要）．Retrieved from http://www.mext.go.jp/b_menu/shingi/chousa/shotou/117/shiryo/__icsFiles/afieldfile/2016/05/24/1368985_7_1.pdf

酒井志延 (2016) リメディアル教育を必要とする学習者を自律的学習者にするための教授法・教材開発「科学研究費助成事業基盤研究（C）（一般）2015 年度研究成果報告書」Retrieved from https://kaken.nii.ac.jp/ja/report/KAKENHI-PROJECT-25370649/25370649seika/

小池生夫 (2010).『企業が求める英語力』English Skills Do Companies Really Need? 朝日出版社.