

The Effectiveness of Dialogue Journal Writing on the Writing Ability of Japanese Learners of English

日本人の英語学習者のライティングスキルに対する ダイアログジャーナルライティングの効果

JOHNSON Stephan & CHENG Benjamin



Key words: 対話ジャーナル執筆、対話ジャーナル、書き込み能力
dialogue journal writing, DJW, Dialogue journals, writing ability

Abstract

This study observed the effects of dialogue journal writing (DJW) as a tool for enhancing writing proficiency of Japanese learners of English. The research compared the effects of DJW and error-corrected feedback (ECF) on two intact classes of first year English learners at a Japanese university. Subjects sat a pre- and post-test in which they were required to write an essay. Scores within groups and between tests were evaluated then compared between groups. Results indicated that learners who adopted DJW did show significant improvements in overall writing ability, however there were no significant improvements in vocabulary usage. Participants in the ECF group showed significant gains in overall writing ability, and some significant changes in vocabulary usage. This research found that DJW proved to be more effective than ECF for improving overall writing ability, but ECF proved to be more effective for improving vocabulary usage among Japanese learners of English.

1. Introduction

Journal writing can be an effective method for improving learners' English ability; Casanave (1993) argues that '[i]n the Japan context, at least, journal writing may constitute the single most beneficial activity for the development of students' confidence and communicative ability in English'. Dialogue journal writing (DJW) is not based on strict methods of teaching, but rather focuses on natural communication between student and teacher such as face-to-face conversations, over a set period of time. This allows the student to express their opinions without the stress and pressure of regular assessed written assignments. Martin, D'Arcy, Newton, and Parker (1976) argue that students feel apprehensive about taking risks because school writing is usually graded and criticised, possibly hindering students' development and their ability to write with confidence. DJW enables students to acquire the language by focusing on meaningful communication; 'the main reason cited for using journals is that they are seen as providing opportunities for authentic, meaningful communication which is focused on the message rather than the form; and that by doing so students will acquire the language' (Duppenhaler, 2004a, p. 2).

From the observations of using DJW in Japanese universities, students improved on their general writing ability. Students learned to correct their mistakes by mirroring the model sentences from the teacher's entries. This did not happen instantly but over a series of journal entries where the mistakes were modelled repeatedly. Furthermore, most students enjoyed communicating via the journals as there were few opportunities for them to converse with their teacher.

1. 1. Dialogue journal writing (DJW)

There are many advocates of dialogue journals (Baskin, 1994; Danielson, 1988; Dooley, 1987; El Koumy, 1998; Liao & Wong, 2010). Baskin (1994) suggests that in order to be effective, teachers need to know and understand the capabilities, needs, and interests of their students as individuals. Teachers can then utilise that knowledge and tailor classes and teaching points accordingly. The use of dialogue journals enables teachers to acquire this knowledge due to the personal nature of journals. A significant benefit to students that resulted from the use of dialogue journals was the improvement in grammar acquisition (Duppenhaler, 2004a; Peyton, 1986). Reading and studying the teacher's entries and having to respond helped improve their grammar knowledge (Baskin, 1994).

Vocabulary usage may also be enhanced through DJW. According to Hamzah, Kafipour, and Abdullah (2009), vocabulary is one of the most important aspects when learning the

meaning of any language. However, they found that English education in universities in Iran was similar to that in Asian universities, in that little emphasis is generally placed on vocabulary teaching. They suggested that dialogue journals provide an opportunity for students to encounter new words. Students can get clear examples of the form of the new words from the teachers' entries, enabling identification of the connection between form and meaning and leading to strong memory connections when attempting to use these new words in their responses.

Vocabulary size is important to learners as it is a good predictor of writing quality (Astika, 1993), and using lexical items effectively is an important factor when writing compositions well (Laufer & Nation, 1995). Bacha (2009) suggested that vocabulary is considered by teachers to be a very important aspect of writing skill, and that 'incoming students need to work on widening their vocabulary repertoire in an academic setting' (p. 379). Grobe (1981) stated that 'what teachers currently perceived as 'good' writing is closely associated with vocabulary diversity' (p. 85).

According to Schmitt and McCarthy (2008), incidental learning, which can occur during activities or exercises that focus on communication, such as DJW, can be an effective, albeit slow (Schmitt, 2000), method for learning large amounts of vocabulary, as learning from context is so important. But how many words are necessary for a language learner to write productively? There are many words that are used in a language but not all of them are useful. Word frequency lists are one measure to gauge the usefulness of words and the range of vocabulary a writer can use (Breeze, 2008).

In the context of Japan, the use of dialogue journals could be beneficial to college Japanese learners of English. DJW can be beneficial to learners who have writing apprehension (Jones, 1991; Peyton, 1990), in that it increases their confidence to write. Mulvey (1999) noted that Japanese learners of English know grammatical rules but cannot speak or communicate in English.

There are possible reasons for the positive effects of DJW. Peyton (1998) suggested that the rapport built between teacher and student, through journals, can motivate students to improve their communicative abilities. As each journal is individualized, teachers can tailor the language appropriately to their students, which can in turn improve communicative ability among the lower level students in a class (El Koumy, 1998). Furthermore, DJW increases opportunities for students to communicate in the target language outside of the classroom (Peyton & Reed, 1990).

Liao and Wong (2010) examined the efficacy of DJW on 10th-grade learners of English in Taiwan, resulting in an improvement in overall writing performance. The study suggested that DJW helped learners to generate more ideas, organise these ideas, and then construct them into compositions of higher quality (Liao & Wong, 2010). However, there was a limitation to the experiment in that it was lacking a control group. Therefore, it was difficult to determine that the improvements in overall writing ability could be solely attributed to DJW. Their paper also collected feedback regarding DJW from the students, with all of them generally having positive responses. More than half of the participants believed that their writing and critical thinking abilities improved, enhancing their confidence in expressing ideas in English.

In the context of Japan, Hirose and Sasaki (2000) studied the effects of teaching metaknowledge and regular journal writing among English students from two Japanese universities. 'Metaknowledge instruction consists of explicitly teaching paragraph elements such as the topic sentence, the body, and concluding sentence, and the types of organisational patterns' (Hirose & Sasaki, 2000, p. 95). Metaknowledge was taught to both groups, but one group was assigned an extra task to write in journals, in the style of a personal diary, outside of regular class time, and this writing was not corrected. The other group received no extra assignments. The other group showed little improvement in their writing ability, whereas the journal-writing group improved significantly.

Duppenthaler (2004a, 2004b) studied the effects of journal writing on a group of 99 second-year students at a private high school in Japan. The one-year experiment conducted over an academic year investigated three types of journal writing feedback for improving writing performance: error correction, positive comments, and meaning-focused feedback (DJW). DJW proved to be the most effective type of feedback. Another important aspect that he discovered was the steady improvement of error-free clauses in journals and in-class writing compositions by participants in the meaning-focused feedback group.

Yoshihara (2008) studied the effects of DJW on writing fluency, which is 'usually measured by the total number of words a writer can produce in a given period of time' (p. 5), of students from a Japanese university. Over two 12-week semesters, the findings concluded that there was no significant improvement in the total number of words produced. This is interesting because many studies have supported the efficacy of DJW for enhancing writing ability in some aspect (Danielson, 1988; Duppenthaler, 2004a, 2004b; El Koumy, 1998; Hirose & Sasaki, 2000; Peyton, 1986; Peyton, 1988; Peyton, 1991; Peyton & Reed,

1990; Staton, 1983), whereas Yoshihara's study is one of the few that did not reveal positive results. However, she noted that the small size of 19 participants could have been a limitation of her study, and may have influenced her results. Furthermore, the students did not sit any tests for the duration of the research. The students wrote in their journals in their own time, and at the end of the experimental period only the journals were analysed. Therefore, one could argue that having the participants sit timed tests before and after the two semesters and then analysing the test papers may have yielded different results. However, according to Duppenhaler (2004b), the skills developed from DJW did not transfer to in-class compositions, and this may also be the case for in-class writing tests, but this is beyond the scope of the present study.

1.2. Error Corrected Feedback (ECF)

This section will outline some of the existing research on the merits of ECF as a tool for marking written compositions, thus providing the rationale for selecting ECF for comparison with DJW in this study. Error corrected feedback is one of the most frequently used methods of student feedback in the classroom (Robb, Shortreed, & Ross, 1986). In addition Applebee believes (1981), error-correction was considered the most important aspect among 80% of foreign language teachers when giving feedback to written compositions. Furthermore, in the ESL classroom, Leki (1991) found that the majority of students viewed error correction, especially in grammatical form, as essential to improving their writing skill. Ferris (2004) asserts that, from a student perspective, because they value error correction so highly the absence could be harmful to their progress, and that learners who receive error correction appreciate the feedback which in turn will motivate them to work on improving their writing, whereas no feedback could result in demotivation and a loss of confidence in their instructors, thereby hindering progress.

There are other researchers that support error correction (Ashwell, 2000; Fathman and Whalley, 1990; Ferris & Roberts, 2001) who all found positive results when comparing compositions of those who received error correction and those who did not. Furthermore, Kepner (1991) found that journal writers who received error corrections on their writings from their teachers made 15% fewer errors than learners who received only comments.

In this paper, error corrected feedback will take the form of correcting all errors in the written compositions.

1.3. Rationale for this study

English teaching in Japan has been a major concern for many years. Japanese students

are taught English for three years in junior high school and three years in senior high school, and a majority of Japanese students study English at least two more years in university. However, until 1999, the Japanese average TOEFL score was approximately 500, and according to the TOEFL Test and Score Data Summary for the years 2001-2002, Japan ranked second from bottom in the computer-based TOEFL and third from bottom in the paper-based TOEFL tests compared with other Asian countries (Yoshida, 2003). In light of this, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been attempting to address this concern since 1989 by setting new objectives to promote higher achievement in English communicative skills among secondary students (Nishino, 2011). However, looking back at the TOEFL results in 2001-2002, MEXT's objectives have clearly not been met.

One method that may promote communicative ability is through writing. In order to enhance the writing skill of Japanese learners, it could be suggested that the issue of grammar instruction in Japanese high schools needs to be addressed. Learners should be made aware of the main purpose of writing (i.e., communication), and should view 'writing as an act of communication with readers to accomplish specific purposes' (Chan, 1986, p. 54).

According to Hirayanagi (1998), Japanese high school teachers place too much importance on grammar rules. It was argued by Braddock, Lloyd-Jones, and Schoer in 1963 that this method of teaching has no beneficial effect on improving writing skill:

'In view of the widespread agreement of research studies based upon many types of students and teachers, the conclusion can be stated in strong and unqualified terms: the teaching of formal grammar has a negligible or, because it usually displaces some instruction and practice in composition, even a harmful effect on improvement in writing' (as cited in Villanueva, 2003, p. 216).

Therefore, DJW may be a solution to improve student's writing abilities through meaningful communication. Moreover, for university students enrolled in an English program, it has been suggested that writing is the most important of the four communication skills because of the requirement to carry out assignments, publish papers, and write dissertations and theses in the second language (L2) (Bagheri & Riasati, 2016).

The present study observed whether DJW is a better tool than error-corrected feedback

(ECF) for enhancing writing skill.

1. 4. Research questions

1. Does DJW significantly improve the overall writing ability of Japanese learners of English compared with ECF?
2. Does DJW significantly improve writers' lexical proficiency compared with ECF?
3. Do learners feel that DJW and ECF are positive tools for learning English?

2. Research methodology

The subjects for this study were two intact first-year classes in a compulsory English communication seminar course at a Japanese university. At the beginning of the semester, the students, aged 18-20 years old, were assigned to classes according to their English ability. There were six classes of approximately 15 students each, with class 1 being the highest and 6 being the lowest English proficiency class. The participants for the present research were classes 3 (intermediate; TOEFL, 420-500) and 6 (upper intermediate, TOEFL, >500).

Class 3 formed the error-corrected feedback group. All students in this class had experience travelling or living overseas, ranging from 1 to 5 weeks. They could deal with most situations likely encountered when travelling in an area where English is spoken. They could produce simple texts on topics which are familiar or of personal interest. Furthermore, they could describe experiences and give brief explanations for opinions.

Class 1 formed the DJW group. Students in this class had experience travelling or living overseas, ranging from 2 weeks to 6 years. These students could interact with a medium to high degree of fluency and spontaneity that makes regular interaction with native speakers possible without any strain. They could produce clear, detailed written texts on a wide range of subjects, and express their viewpoints on topical issues.

The institution would not allow classes to be divided or mixed with other classes, so unfortunately it was only possible to have groups of different proficiencies. This may not be ideal, but it could be suggested that the tools examined in this research to enhance writing proficiency should be effective regardless of ability, and that observing changes within the groups followed by a comparison of these changes between the groups should be sufficient.

2. 1. Methodology

The duration of the study was 14 weeks. Prior to the experiment, the participants took a 40-minute essay writing test, which was marked by two native English teachers (the researcher and the participants' teacher) using a modified version of Hemmati and Soltanpour's (2012) analytic scoring scale measuring five criteria: content, organisation, vocabulary choice, grammar, and spelling and punctuation (see Appendix A). Analytical scoring allows teachers to provide detailed information to writers as to what characteristics of writing are important, and it also helps scorers to assign marks to subjective assessments (Perlman, 2003). The final score was the average of the two teachers' scores. The topics for the pre- and post-test essays were chosen so that the participants did not require previous knowledge of the subject that may affect their writing in order to complete the test (Duppenthaler, 2004b). The learners were required to write approximately 250 words.

The length of written compositions may be a factor when grading. 'It can be said that the lengths of texts are important in assessments carried out by raters because raters generally gave lower rating to short texts. However, this does not mean that all long texts were scored higher' (Beyreli & Ari, 2009, p. 117). Therefore, for the present research, the minimum number of words for the pre- and post-test essays was set at 250, based on the word limits set for writing test questions in the Official TOEFL iBT Tests with Audio, International Edition (Educational Testing Service, 2013).

All papers were graded by the two native teachers and then securely stored for later use. The ECF group wrote a short paragraph about a topic given by the researcher once a week for 15 minutes at the end of class during the experimental period. Errors were corrected and then given back to the students. Learners were not required to revise their compositions. The students wrote a short paragraph about a different topic every week for the duration of the project. The DJW group were given notebooks to write their journal entries. During the experimental period, once a week during the last 15 minutes of class was allocated to writing in their dialogue journals. After each entry, the journals were collected, a response was written, and then the journals were returned for a follow-up reply. This continued for the duration of the study. The dialogue journals and the ECF group's weekly compositions were not analysed at any stage of the project.

At the end of the 14-week experimental period, all participants sat a final 40-minute essay writing test. The papers were marked by the same two native English teachers who scored

the pre-test, with the final score being the average of the two raters. The pre- and post-test scores were then analysed and compared to search for any significant improvements between the groups.

2.2. Data collection

Prior to the commencement of the project, learners participated in a question and answer session regarding the project. The participants were given a detailed explanation of the project before completing a consent form (see Appendix B). This session was necessary so that the participants could receive full disclosure about the experiment and understand what their participation entailed and that the experiment was voluntary and they could withdraw from the study at any time without reason or repercussions. As suggested by Mackey and Gass (2011), the participants were informed of the procedures and aims of the research as outlined above in the methodology. Moreover, the steps taken to ensure their data and identities were kept confidential were provided.

During this same session, participants completed a questionnaire (see Appendix C). This was used to gather more information related to their English studies such as extra English classes or time spent living overseas. The same questions were answered at the end of the project in case there were any changes that may have affected their scores between tests.

All tests and questionnaires were carried out in the participants' regular classrooms. Only the two raters had access to the papers. Once the raters collected the test papers, the students were assigned numbers to safeguard participants' identities before grading. All papers were then kept and stored securely by the researcher for the duration of the project. The scores were entered into a password-protected computer that only the researcher had access to.

At the end of the experiment, the participants answered an anonymous Likert questionnaire (see Appendix D) regarding their thoughts and opinions about the project, which was required by the university. All data collected from the research will be retained for the duration of the project and kept by the researcher for seven years from the publication of the results. At the end of this period, the data will be destroyed.

3. Overall writing ability

Odell (1981) defines competence in writing as when a writer has the ability to convey their message using language in a way that they wish by producing content that is

appropriate for the intended audience and purpose. Charney (1984) states that quantitative methods are not a suitable means of assessing written compositions because they only measure rigid, 'plain' constructions as opposed to a writer's ability to form writing samples that show a high quality of writing style that does not follow general writing conventions. He argues that qualitative methods are more valid for measuring the quality of writing skill. Therefore, the present research utilized a subjective method of assessment.

3. 1. Subjective scoring

Perkins (1983) believes that the quality of writing is connected with the communication of meaning, and that objective measures do not show a writer's ability to communicate. Therefore, Perkins suggested that objective measures have little value when it comes to assessing communication aspects in writing. Furthermore, objective measures do not account for many factors that make up a well-written composition. These include areas such as the use of idioms, presentation and development of ideas in an organized fashion, sentence complexity and variety (Polio, 1997), and the relevance of arguments to name a few. Perlman (2003) states that it is necessary to include some subjective judgments when scoring the performance of students' writing quality.

The two kinds of subjective scoring that are most commonly used are holistic and analytical scoring. Holistic scoring entails the rater evaluating a piece of work based on an overall impression that is usually in the form of a letter grade, a percentage, or a number on a scale that corresponds to a given criteria. This type of scoring is useful when scoring classroom essays as it is efficient and saves time and money. Analytic evaluation requires raters to award scores for specific components of writing using a rubric detailing the corresponding criteria. The raters first give a score to each element, such as organization, content, grammar, and mechanics, and then combine them to give a total score. This form of evaluation has been preferred to holistic scales as it has been indicated that writers may perform differently in each writing aspect, thereby making analytic rubrics necessary when evaluating specific features of writing ability (Bacha, 2001).

3. 2. Analytic scoring

Bacha (2001) carried out a study on final exam essays written in Arab students in an EFL program and discovered that analytic scoring was beneficial in informing teachers of their students' essay proficiency. The findings of the study showed that by being criteria specific, such as having scales to measure structural patterns or lexical features, analytic scoring scales are well suited when evaluating different aspects of learners' writing ability.

Jonsson (2007) reviewed 75 empirical studies on rubrics, concluding that the use of analytic rubrics was a reliable tool for assessment. The study noted that a benefit of using scoring rubrics was that they enhanced the consistency of scoring between different raters.

It may also be argued that using analytic rubrics are more valid and reliable than not adopting one for evaluating writing skill. Rezaei and Lovorn (2010) conducted a study on the reliability and validity of using rubrics for assessing written compositions. Participants graded two sample essays prepared by the researchers, one with a rubric and one without. The findings showed that the participants' ratings were highly influenced by mechanical aspects (such as grammar, spelling, punctuation, and sentence structure) of the compositions as opposed to content.

Mikan argues that grammatical and vocabulary features are difficult to determine, and participants may not be graded according to their writing ability (cited in Bagheri & Riasati, 2016). For example, an essay that contained sophisticated ideas but incorrect morphemes may be penalised and therefore receive a low grade, even though the quality of the writing was considered to be high. Hence, grading a combination of writing attributes, such as content, structure, lexical choice, and grammatical features may prove to be a more appropriate means of evaluation. Moreover, analytic scoring has been used extensively when measuring overall writing ability (Engber, 1995).

After considering the literature, the use of an analytic scoring rubric, for assessing the overall writing performance of student essays was utilised when evaluating the participants' pre- and post-test essays.

3. 3. Data analysis

For the present study, there were two raters because two raters are sufficient for producing acceptable inter-rater reliability levels, and 'intra-rater reliability might not in fact be a major concern when raters are supported by a rubric' (Jonsson, 2007, p. 133). Jonsson's research also reported that most studies investigating intra-rater reliability found that when raters used rubrics for assessment their consistency was sufficient.

3. 3. 1. Assumptions

A Shapiro-Wilk test was carried out on the test scores to check for normal distributions. The pre-test scores of the DJW group ($p = 0.36$, not significant at $p < 0.05$) were normally distributed. The DJW group's post-test scores also had a normal distribution ($p = 0.39$, not

significant at $p < 0.05$). A Levene's test revealed that the variances were not equal ($p = 0.04$, significant at $p < 0.05$). The results from these tests showed that the DJW data were non-parametric, and therefore the Wilcoxon signed-rank test was used for analysing the pre- and post-test scores.

The pre-test scores from the ECF group were normally distributed ($p = 0.96$, not significant at $p < 0.05$). However, the post-test scores were not ($p = 0.04$, significant at $p < 0.05$) resulting in rejecting the null hypothesis that the data were normally distributed. As a result, the Wilcoxon signed-rank test was also used for analysing the pre- and post-test scores for the ECF group.

3.3.2. Results

Table 1. Wilcoxon signed-rank test for the rubric scores between groups

		Median**	<i>W</i>	<i>Z</i>	<i>p</i>	<i>r</i>
ECF n = 15	Pre-test	14.50	12.50	-2.70	<0.01*	-0.49
	Post-test	15.50				
DJW n = 16	Pre-test	18.75	1.50	-3.32	<0.01*	-0.59
	Post-test	22.25				

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback.

* $p < 0.05$; **Maximum possible score, 25.

Table 1 shows that the DJW group improved significantly from the pre-test (median, 18.75) to the post test (median, 22.25; $W(25) = 1.5$; $p = 0.0009$, significant at $p < 0.05$ with a strong effect; $r = 0.59$). The ECF group also displayed significant improvements between tests. The medians of the pre- and post-test scores in the ECF group were 14.5 and 15.5, respectively ($W(25) = 12.5$; $p = 0.0069$, significant at $p < .05$ with a medium to strong effect; $r = 0.49$).

The overall analytic scores from the pre- and post-tests between groups (see Appendix F) were also observed, with some interesting points coming to light. From the ECF group, 3 of the 15 participants displayed a negative change in scores, ranging from -2% to -6%, compared with only 1 of the 16 participants from the DJW group (with a change of -2%). The largest improvement from the ECF group was a 20% increase (one participant) between tests, compared with a 32% increase (two participants) from the DJW group. Furthermore, 53% of the participants in the ECF group improved their scores by 1–10% between tests, whereas 12% of those from the DJW group had. However, 56% of the participants in the DJW group improved their scores by 11–20%, compared with 13% of those in the ECF group. Finally, 19% of the participants in the DJW group improved their

scores between tests by at least 21%, compared with none from the ECF group, indicating that the DJW group made larger gains between tests.

4. Vocabulary usage

The present research observed the percentage of word families used from the 2,000 most frequent word families (K1–K2 frequency bands) and those in the mid-frequency range and above (K3–K25 frequency bands) used in the pre- and post-test compositions. Bradford (1979) carried out a study of short-term changes in the writing skills of EFL students and concluded that one of the most significant changes was the size of vocabulary. Leki and Carson (1994) surveyed university undergraduate students, asking what they considered were important skills for their English for Academic Purposes writing courses. The results showed that vocabulary was deemed the most important. Leki and Carson concluded that students were concerned with having the ability to use specific lexical items effectively for academic writing. Moreover, Arnaud (1984) notes the phenomenon that the quality of an essay is lowered in some L2 essays that contain very few lexical items and are repeated frequently.

Llach and Gallego (2009) discovered that the correlation between essay quality and receptive vocabulary of Spanish primary school EFL learners, although not very high, was significant. Their research stated that having knowledge and possession of a large amount of vocabulary contributed to higher-quality written compositions. For example, a high-quality essay would consist of the use of lower frequency words, a variety of different words, more non-repeated words, and words that have not been used by other peers. However, they did note that even though compositions that contain high lexical richness will receive a higher score, they do not have particularly strong correlations, and other factors such as mechanical or content structures may also influence the evaluation of a composition.

4. 1. Data analysis

According to Nation and Waring (1997), knowledge of approximately 2,000 to 3,000 of the most frequent word families is sufficient for productive writing. Nation (2000) reported that most academic texts contain 87% of words from the 2,000 most frequent word families, and that knowing these is crucial for writing effectively. Schmitt and Schmitt (2014) recommend that university students should be knowledgeable of word families from the middle frequency range (the most frequent 3000–9000 word families), as many university students who possessed a vocabulary size of around 3,000 word families

struggled with their courses due to difficulties in reading the required university academic texts.

Nation's VocabProfiler (<http://www.lextutor.ca/vp/eng/>) was used to measure vocabulary richness because it 'has been shown to be a reliable and valid measure of lexical use in writing' (Laufer and Nation, 1995). Moreover, it can be used to identify low proficiency learners who may have problems in their study programs, thus allowing institutions to address their weaknesses by teaching the required vocabulary effectively (Morris and Cobb, 2004).

The VocabProfiler package shows the variety of vocabulary a writer uses at different frequency levels, and therefore can identify the lexical richness of a composition by observing the percentage of word families within a text. In the present study, the BNC-COCA-25 VocabProfiler version was used, which allows the K1–K25 frequency bands to be analyzed in order to observe whether the participants produced more or less word families from the different frequency bands.

4. 1. 1. Assumptions

A Shapiro–Wilk test was carried out on the percentage of the first 2,000 most frequent word families used (K1–K2 frequency band) to check for normal distributions. The pre-test data from the ECF group were found to not have normal distributions ($p < 0.05$); however, the post-test data ($p > 0.05$) was normally distributed. The same test was carried out on the percentage of vocabulary used above the 2,000 most frequent word families (K3–K25 frequency band), revealing similar results, wherein the pre-test ($p < 0.05$) did not have normally distributed data, but the post-test data ($p > 0.05$) did.

After testing the DJW group, it was found that the pre-test data for the percentage of the 2,000 frequency band were not normally distributed ($p = 0.04$, significant at $p < 0.05$), but the post-test data were ($p = 0.29$, not significant at $p < 0.05$). The pre-test data for the K3–K25 frequency band were not normally distributed ($p = 0.04$), but the post-test data were ($p = 0.32$). Therefore, as a result of the Shapiro–Wilk test findings, the Wilcoxon signed-rank test was used for analyzing vocabulary usage of the groups.

4. 1. 2. Results

Table 2. Wilcoxon signed-rank test for percentage of the K1–K2 frequency bands (2,000 most frequent word families)

		Median*	<i>W</i>	<i>Z</i>	<i>p</i>	<i>r</i>
ECF n = 15	Pre-test	94.44				
	Post-test	95.45	26.00	-1.93	0.05	-0.35
DJW n = 16	Pre-test	93.80				
	Post-test	94.40	33.00	-1.81	0.07	-0.32

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback.

*Median percentage of K1-K2 word families in writers' texts.

Table 2 shows that participants in the ECF group did not make significant changes from the 2,000 most frequent word families in the post test. The medians of the pre- and post-test percentages in the ECF group were 94.44 and 95.45, respectively ($W(25) = 26$; $p = 0.0536$, not significant at $p < 0.05$; $r = 0.35$). The DJW group showed similar results (pre-test median, 93.80; post-test median, 94.40; $W(29) = 33$; $p = 0.07$, not significant at $p < 0.05$; $r = 0.32$).

Table 3. Wilcoxon signed-rank test for percentage of the K3–K25 frequency bands (2,001 and above most frequent word families)

		Median**	<i>W</i>	<i>Z</i>	<i>p</i>	<i>r</i>
ECF n = 15	Pre-test	5.56				
	Post-test	4.55	21.00	-1.98	0.05*	-0.36
DJW n = 16	Pre-test	6.19				
	Post-test	5.61	33.00	-1.81	0.07	-0.32

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback.

* $p < 0.05$; **Median percentage of K3–K25 word families in writers' texts.

Table 3 shows that participants in the ECF group showed a more significant change in producing word families from the K3-K25 frequency band in the post-test (median, 4.55) than in the pre-test (median, 5.56; $W(21) = 21$; $p = 0.048$, significant at $p < 0.05$; $r = 0.36$).

5. Questionnaire results and findings

This summary provides a brief description of the responses from the pre- and post-experiment questionnaires regarding the extra information of the participants' English studies. It was found that there was no change during the experimental period. Furthermore, this section provides student feedback and comments from the Likert questionnaire. These data (see Appendix G for tables of Likert questionnaire responses)

were aggregated from a questionnaire (comprising ten 5-point Likert-scaled items) that students were asked to complete at the end of the course. The responses “somewhat agree” and “strongly agree” have been consolidated into one category in the summary report below for the sake of simplicity.

5. 1. ECF group, n = 15

93% of respondents agreed that the course content matched the course description on the syllabus. All agreed that the in-class writing assignments were helpful. Ninety-three percent agreed that the in-class assignments were manageable. All agreed that class size was ideal. 80% agreed that the English proficiency among students was not a problem. 86% agreed that the in-class writing had inspired them to write more than before.

5. 2. DJW group, n = 16

Seventy-five percent of respondents agreed that the course content matched the course description in the syllabus. Eighty-seven and five tenths percent agreed that the in-class writing assignments were helpful. Ninety-three percent agreed that the in-class assignments were manageable. All agreed that class size was ideal. Ninety-three percent agreed that the English proficiency among students was not a problem. Eighty-one percent agreed that the in-class writing had inspired them to write more than before.

5. 3. 1. ECF group, n = 15

There were nine feedback comments from this group, all positive. These comments were corrected for ease of reading and are as follows: ‘writing in class was a good experience to improve my writing skills’, and ‘I think this class was worthwhile’. There were students who found the writing useful: ‘I think the experience of writing and presentation in this class will be useful for me’. There was one participant who felt that their vocabulary required improvement: ‘I also felt that I have to improve my vocabulary’.

5. 3. 2. DJW group, n = 16

There were only four comments from this group, but all were positive. One student wrote, ‘This class gave me a lot of opportunities to write reports in English so I feel that I can improve my English skills’. Another wrote, ‘This class really inspired me’. From these comments, it could be concluded that learners who use ECF and DJW feel that these tools are beneficial for improving English skill.

6. Discussion of results

6. 1. Vocabulary usage

One could argue that these results should be taken lightly due to the method of collecting vocabulary data used in this research. Laufer (1998) stated that the use of free composition writing may not be an accurate tool to distinguish one's vocabulary knowledge, as learners may not use the infrequent lexical items they know when left to their own free will. He also suggested that learners may have to learn a large amount of lexical items passively before some of them can be used freely. Moreover, many L2 learners have been taught English through methodologies that place emphasis on everyday communication; therefore, it should not be surprising that their written compositions resemble spoken English containing basic structures that may be less appropriate for academic writing (Breeze, 2008).

According to Folse (2006), testing vocabulary usage through writing exercises, such as the ones carried out in the present research, may not be the most effective, as vocabulary could be measured in other ways that do not require written compositions. He states that writing original sentences is not an effective or efficient method for students to grow and retain L2 vocabulary. He concluded that so-called "fill-in-the-blank" exercises would be more suitable because of the time involved in writing original sentences. By creating original sentences, a large amount of work would be required from the learner to perform tasks such as finding the word in a dictionary, deciding if the word can be used in a way that they would like, producing a suitable sentence, and then deciding if the sentence is correct. Folse stated that by learning new vocabulary through this method, sentences may contain errors, specifically those involving usage and collocation. For the teacher, correction would be time consuming and it would prove difficult to re-word so that the writer's original meaning is maintained. Folse asserted that if the time spent produced solid results for vocabulary retention, then the efforts of students and teachers would be warranted.

It has been contended that teachers do not teach or use mid-frequency vocabulary (the most frequent 3,000–9,000 word families) in the classroom, instead focusing on high-frequency vocabulary (Schmitt & Schmitt, 2014). This creates a pedagogical requirement that needs to be addressed, considering the importance and benefits of using mid-frequency vocabulary. Learners who are unable to access and use lower-frequency vocabulary at an academic level will have more difficulties as the demands of higher-

quality writing increase (Morris & Cobb, 2004).

6. 2. Observations

From the present study, it was observed that some students wrote with more speed and fluency as the course progressed, and the number of words produced between tests increased (see Appendix J). A Shapiro–Wilk test ($p > 0.05$) carried out on all data for the number of words written between tests and for both groups followed by a Levine’s test ($p > 0.05$) concluded that all data were parametric. A dependent means t -test revealed that the ECF group wrote significantly longer compositions ($p = 0.00015$, significant at $p < 0.05$; Table 4). However, the DJW group did not ($p = 0.23$, not significant at $p < 0.05$).

Table 4. Dependent means t -test for total number of words produced

		Mean	SD	t	df	p	r
ECF n = 15	Pre-test	154.33	50.63	5.15	14.00	<0.01*	0.83
	Post-test	252.13	58.56				
DJW n = 16	Pre-test	255.38	71.26	1.26	15.00	0.23	0.32
	Post-test	277.13	48.05				

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback; SD, standard deviation.

* $p < 0.05$.

It has often been claimed that an increase in writing fluency is one of the benefits of DJW, and that DJW enables learners to become confident in taking risks with their writing and focusing less on grammatical accuracy, thereby enhancing fluency (Casanave, 1994). However, the results from this research contradict this and are in line with Yoshihara (2008), who concluded that Japanese learners of English do not statistically increase the number of words they can produce in an allotted period by writing in journals.

A possible explanation for this is that English ability may have contributed to the outcome. As mentioned previously, the ECF group were the intermediate-level class and the DJW group were the advanced-level class. It could be suggested that gains of lower-level learners are readily noticeable, but gains for advanced levels are not. According to Casanave (1994), it is difficult to notice improvements over time of advanced-level learners.

7. Conclusions

This research sought to contribute to the few empirical studies on the effects of DJW as a tool for improving writing ability in the context of Japan. Previous research on DJW has

generally concluded that this tool improves writing in many aspects. These range from areas such as grammar acquisition, vocabulary usage, and fluency. In the context of Japan, there has been little research to support these claims. The current study observed the effects of DJW and ECF on first-year students in a Japanese university and concluded that DJW is more effective for improving overall writing ability than ECF. However, ECF was found to be a better tool for enhancing lexical proficiency in terms of vocabulary usage.

First, the study revealed that DJW did not improve vocabulary usage. However, DJW did significantly improve participants overall writing ability in a subjective sense, and learners felt that it was a positive tool for improving their English. Second, ECF significantly improved overall writing ability. In terms of lexical competency, writers made some improvements, in that they significantly increased their usage of less-frequent vocabulary (vocabulary above the most frequent 2,000 word families), but they did not decrease their usage of the most frequent vocabulary (the 2,000 most frequent word families). Consequently, learners felt that ECF had a positive impact on their English language learning.

This research did not represent a true reflection of the population due to the small sample sizes (15 participants in the DJW group and 16 in the ECF group). Therefore, a larger sample size of more than 40 subjects per group may yield more reliable results. Furthermore, although the results from the current research revealed that DJW did not significantly improve learners' proficiency in vocabulary usage in participants' written compositions, it will be interesting to observe the effects of comparing ECF with DJW between groups with similar English ability in the context of Japan.

Bibliography

- Applebee, A. N. (1981). *Writing in the secondary school* (NCTE Research Rep. No. 21). Urbana, IL: National Council of Teachers of English.
- Arnaud, P. J. L. (1984) The lexical richness of L2 written productions and the validity of vocabulary tests. (ERIC Document Reproduction Service No. ED 275 164/FL 016 068).
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9, 227–258.
- Astika, G. G. (1993). Analytical assessment of foreign students' writing. *RELC Journal*, 24, 61-72.
- Bacha, N. (2001). Writing evaluation: what can analytic versus holistic essay scoring tell us? *System*, 29, 371-383.
- Bagheri, M. S. & Riasati, M. J. (2016). EFL graduate students' IELTS writing problems and students' and teachers' beliefs and suggestions regarding writing skill improvement.

- Journal of Language Teaching and Research*, 7(1), 198-209.
- Baskin, R. S. (1994). Student feedback on dialogue journals. (ERIC Document Reproduction Service No. ED 375 627 / FL 022 459).
- Beyreli, L. & Ari, G. (2009). The use of analytic rubric in the assessment of writing performance. Inter-rater concordance study. *Educational Sciences: Theory and Practice*, 9(1), 105-125.
- Bradford, A. (1979). Short-term changes in ESL composition. In Yorio C., Perkins, K., & Schachter, J. (Eds), *On TESOL '79* (pp. 330-342). Washington, D. C.: TESOL.
- Breeze, R. (2008). Researching simplicity and sophistication in student writing. *IJES*, 8(1), 51-66.
- Casanave, P. C. (1993). 'Introduction', in Casanave, P. C. (ed.) *Journal Writing Pedagogical Perspectives*. Fujisawa: Keio University, Institute of Languages and Communication, pp. 1-5
- Casanave, P. C. (1994). Language development in students' journals. *Journal of Second Language Writing*, 3(3), 179-201.
- Casanave, P. C. (2004). *Controversies in second language writing: Dilemmas and decisions in research and instruction*. Ann Arbor, MI: The University of Michigan Press.
- Chan, M. M. (1986). Teaching writing as a process of communication at the tertiary level. *JALT Journal*, 8(1), 53-70.
- Charney, D. (1984). The validity of using holistic scoring to evaluate writing: a critical overview. *Research in the Teaching of English*, 18(1), 65-81.
- Danielson, K.E. (1988). *Dialogue Journals: Writing as Conversation*. Bloomington, IN: Phi Delta Kappa Educational Foundation (ERIC Document Reproduction Service No. ED 293 144 / CS 211 136).
- Dooley, M. S. (1987). Dialogue journals: facilitating the reading-writing connection with Native American students. (ERIC Document Reproduction Service No. ED 292 118 / CS 211 095).
- Duppenthaler, P. (2004a). The effect of three types of feedback on the journal writing of EFL Japanese students. *JACET Bulletin*, 38, 1-17.
- Duppenthaler, P. (2004b). Journal writing and the question of transfer of skills to other types of writing. *JALT Journal*, 26(2), 171-188.
- Educational Testing Service. (2013). *Official TOEFL iBT tests with audio international edition*. Asia[[Singapore]: McGraw-Hill Education.
- El-Koumy, A. S. A. (1998). Effect of dialogue journal writing on EFL students' speaking skill. (ERIC Document Reproduction Service No. ED 424 772 / FL 025 572).
- Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press

- Ferris, D.R. (2004) The "Grammar Correction" Debate in L2 Writing: Where are we, and where do we go from here? (and what do we do in the meantime ...?) Dana R. Ferris* *Journal of Second Language Writing* 13 (2004) 49–62
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10, 161–184.
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273-294.
- Grobe, C. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 15, 75-85.
- Hamzah, M. S. G., Kafipour, R., & Abdullah, S. K. (2009). Vocabulary learning strategies of Iranian undergraduate EFL students and its relation to their vocabulary size. *European Journal of Social Sciences*, 11(1), 39-50.
- Hemmati, F., & Soltanpour, F. (2012). A comparison of the effects of reflective learning portfolios and dialogue journal writing on Iranian EFL learner's accuracy in writing performance. *English Language Teaching*, 5(11), 16-28.
- Hirayanagi, Y. (1998). Writing to improve analytical and organizational skills. *The Language Teacher*, 22(12), 21-23.
- Hirose, K., & Sasaki, M. (2000). Effects of teaching metaknowledge and journal writing on Japanese university students' EFL writing. *JALT Journal*, 22 (1), 94-113.
- Jones, P. (1991). What are dialogue journals? In Peyton, J. K., & Staton, J. (Eds.), *Writing our lives: Reflections on dialogue journal writing with adults learning English* (pp. 3-10). Englewood Cliffs, NJ: Prentice-Hall.
- Jonsson, A. (2007). The use of scoring rubrics: reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, 75, 305–313.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*, 19(2), 255-271.
- Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals*, 24, 203–218
- Leki, I., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the discipline. *TESOL Quarterly*, 28, 81-101.
- Liao, M. T., & Wong, C.T. (2010). Effects of dialogue journals on L2 students' writing fluency, reflections, anxiety, and motivation. *Reflections on English Language Teaching*, 9(2), 139-170.
- Llach, M. P., & Gallego, M. T. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Journal of the Spanish Association of Anglo-American Studies*, 31(1), 129-147.

- Mackey, A., & Gass, S. M. (2011). *Second language research: methodology and design*. New York: Routledge.
- Martin, N., D'Arcy, P., Newton, B., & Parker, R. (1976). *Writing and learning across the curriculum 11-16. Schools Council Project*. Upper Montclair, NJ: Boynton/Cook Publishers.
- Morris, L., & Cobb, T. (2004). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System*, 32, 75-87.
- Mulvey, B. (1999). A myth of influence: Japanese university entrance exams and their effect on junior and senior high school reading pedagogy. *JALT*, 21(1), 125-142.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In Schmitt, N., & McCarthy, M. (eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Nishino, T. (2011). Japanese high school teacher's beliefs and practices regarding communicative language teaching. *JALT Journal*, 33(2), 131-156.
- Odell, L. (1981). 'Defining and Assessing Competence in Writing English', in Cooper, C. R. (ed.) *The Nature and Measurement of Competency in English*. Urbana, Illinois: National Council of Teachers of English, pp. 95-138.
- Perlman, C. C. (2003). Performance assessment: designing appropriate performance tasks and scoring rubrics. (ERIC Document Reproduction Service No. ED 480 070 / CG 032 643).
- Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly*, 17(4), 122-142.
- Peyton, J. K. (1990). Dialogue journal writing and the acquisition of English grammatical morphology. In Peyton, J. K. (ed.), *Students and teachers writing together: perspectives on journal writing* (pp. 67-97). Alexandria, VA: Teachers of English to Speakers of Other Languages.
- Peyton, J. K. (1988). Mutual conversations: written dialogue as a basis for building student-teacher rapport. In Staton, J., Shuy, R. W., Peyton, J. K., & Reed, L. (eds.), *Dialogue journal communication: classroom, linguistic, social and cognitive views* (pp. 183-201). Norwood, NJ: Ablex Publishing Corporation
- Peyton, J. K. (1986) Dialogue journal writing and the acquisition of English grammatical morphology. (ERIC Document Reproduction Service No. ED 276 257 / FL 016 209).
- Peyton, J. K., & Reed, L. (1990). *Dialogue journal writing with nonnative speakers: a handbook for teachers*. Alexandria, VA: TESOL.
- Peyton, J. K., & Staton, J. (eds) (1991). *Writing our lives: reflections on dialogue journal writing with adults learning English*. Englewood Cliffs, NJ: Prentice Hall Regents.
- Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 47(1), 101-143.
- Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15, 18-39.
- Robb, T., Shortreed, I., & Ross, S. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20(1), 83-95.

- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., & McCarthy, M. (2008). *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47, 484-503. DOI: 10.1017/S0261444812000018
- Staton, J. (1983). Dialogue journals: a new tool for teaching communication. (ERIC Document Reproduction Service No. ED 227 701 / FL 013 572).
- Villanueva, V. (2003). Cross-talk in comp theory: a reader. Second edition, revised and updated. (ERIC Document Reproduction Service No. ED 474 209 / CS 511 853).
- Yoshihara, R. (2008). The bridge between students and teachers: the effect of dialogue journal writing. *The Language Teacher*, 32(11), 3-7.
- Yoshida, K. (2003). Language education policy in Japan – the problem of espoused objectives versus practice. *Modern Language Journal*, 87(2), 291-291.

Appendices

Appendix A – Analytic scoring scale

	5 Advanced– High	4 Advanced–Low	3 Intermediate– High	2 Intermediate– Low	1 Novice
Content * Logical development of ideas * Main ideas, supporting ideas, and examples	Effectively addresses the topic and task, using clearly appropriate explanations, examples, and details	Addresses the topic and task with using appropriate explanations, examples, and details	Addresses the topic and task using somewhat developed explanations and details	Limited development in response to the topic and task using inappropriate explanations, examples, and details	Questionable responsiveness to the topic and task using no detail or irrelevant explanations
Organization * The sequence of introduction, body, and conclusion	Well organized and cohesive devices effectively used	Fairly well organized and cohesive devices adequately used	Loosely organized and incomplete sequencing; cohesive devices may be absent or misused.	Ideas are disconnected and lack of logical sequencing; inadequate order of ideas	No organization and no use of cohesive devices
Language in use * Choice of vocabulary	Appropriate choice of words and use of idioms	Relatively appropriate choice of words and use of idioms	Adequate choice of words but some misuse of vocabulary or idioms	Limited range of vocabulary, confused use of words and idioms	Very limited vocabulary, very poor knowledge of idioms
Grammar *Sentence-level structure	No errors, full control of syntactic variety	Almost no errors, good control of syntactic variety	Some errors, fair control of syntactic variety	Many errors, poor control of syntactic variety	Severe and persistent errors, no control of syntactic variety
Mechanics *Punctuation *Spelling *Capitalization *Indentation	Mastery of spelling and punctuation	Few errors in spelling and punctuation	Fair number of spelling and punctuation errors	Frequent errors in spelling and punctuation	No control over spelling and punctuation

Appendix B – Consent form

The Effectiveness of Dialogue Journals Consent Form

This study will investigate the effects of using dialogue journals as a tool for improving writing skills among L2 students in a Japanese university. The duration of the study will be 14 weeks.

Before the experiment commences, all participants will be given approximately five minutes to answer questions in the presence of the researcher so that any queries can be addressed. The questions will be used to gather background information regarding students' habits towards studying English such as time spent living abroad or extra English classes outside of school. At the end of the study all students will answer the same questions.

- YES NO I confirm that the purpose of the study has been explained and that I have understood it.
- YES NO I have had the opportunity to ask questions and they have been successfully answered.
- YES NO I understand that my participation in this study is voluntary and that I am free to withdraw from the study at any time, without giving a reason and without consequence.
- YES NO I understand that all data are anonymous and that there will not be any connection between the personal information provided and the data.
- YES NO I understand that there are no known risks or hazards associated with participating in this study.
- YES NO I confirm that I have read and understood the above information and that I agree to participate in this study.

Participant's Name AND Signature: _____

Researcher's Signature: _____

Date: _____

Appendix C – Pre-experiment questionnaire

Your English Experience

Please take a few minutes to complete this questionnaire and give it to the instructor.

1. Age

2. Where and how long have you studied English? (Check the following.)

Elementary School years

Junior High school years

High school years

University years

Institutions years

3. Have you ever traveled to or lived in an English-speaking country?

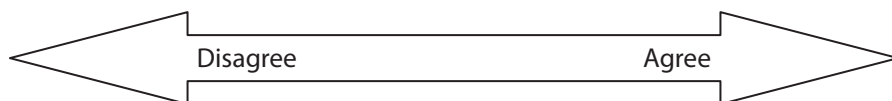
a) Where? (Please specify)

b) For how long? (Please specify)

Appendix D – End of experiment Likert questionnaire

How are we doing?

Please take a few minutes to complete this questionnaire and give it to the instructor.



--2 Somewhat disagree -1 0 N/A +1 Somewhat agree ++2 Strongly disagree
Strongly agree

Please circle to most appropriate response to each of the statements below:

1. The course content matched the course description.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

2. The in-class writing assignments were helpful.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

3. The in-class writing was manageable.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

4. The size of the class was ideal.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

5. The English proficiency among students was not a problem.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

6. This in-class writing has inspired me to write more than before.

--2	-1	0	+1	++2
Strongly disagree	Somewhat disagree	N/A	Somewhat agree	Strongly agree

7. Age

8. Where and how long have you studied English? (Check the following.)

Elementary School years

Junior High school years

High school years

University years

Institutions years

9. Have you ever traveled to or lived in an English-speaking country?

a) Where? (Please specify)

b) For how long? (Please specify)

Note: Feel free to add any comments you might have on this side of this paper. Thanks!

Notes - (comprised of ten, 5-point Likert items)

Appendix E – Percentage of error-free T-units and mean T-unit length

Student	ECF group				Student	DJW group			
	Pre-test		Post-test			Pre-test		Post-test	
	P of E	TL	P of E	TL		P of E	TL	P of E	TL
1	31.8	12.273	46.7	16.73	1	47.62	14.24	53.3	17.133
2	33.3	11.333	69.6	13.96	2	47.62	12.29	62.1	10.345
3	71.4	10.143	86.4	11.09	3	41.67	13	25	12.125
4	56.3	11.563	66.7	16.33	4	27.27	17.73	57.1	14.143
5	71.4	11.857	71.4	9.286	5	47.06	12.76	53.8	11.115
6	57.1	12.357	50	16.69	6	70.6	14.12	52.9	16.412
7	28.6	10.357	31.3	12.13	7	81.25	14.69	72.2	14.222
8	50	11.5	57.9	20.37	8	85.71	24.14	86.7	20.733
9	72.7	11.818	50	14.25	9	50	15.79	28.6	16.929
10	27.8	10.611	38.5	12.77	10	65.52	14.1	68.4	15.053
11	33.3	10.333	50	13.79	11	47.37	14.21	44.4	15.667
12	60	10.267	76.5	15.12	12	42.86	15.48	56	13.52
13	35.3	11.353	42.3	12.12	13	69.57	13.17	47.4	10.316
14	57.1	9.8095	47.4	13.58	14	56.25	15.63	59.3	13.519
15	50	11.6	80	13.67	15	56.25	16.13	57.1	15.333
					16	100	19.17	53.8	17.154

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback; P of E, percentage of error-free T-units within the text; TL, mean T-unit length.

Appendix F – Comparison of the pre- and post-test analytic scores

ECF group				DJW group			
Student	Pre-test	Post-test	Change %	Student	Pre-test	Post-test	Change %
1	16	14.5	-6	1	21.5	22	2
2	15	16.5	6	2	16.5	20	14
3	9.5	14.5	20	3	17.5	22.5	20
4	14.5	15.5	4	4	14.5	20.5	24
5	12	14.5	10	5	17.5	21	14
6	16.5	15.5	-4	6	20	20	0
7	13	14.5	6	7	15.5	23.5	32
8	15.5	19.5	16	8	20	24.5	18
9	15.5	16.5	4	9	15	23	32
10	13.5	13	-2	10	22	21.5	-2
11	11.5	14	10	11	18.5	21.5	12
12	12.5	15.5	12	12	20.5	23.5	12
13	17.5	19	6	13	15.5	19.5	16
14	14.5	15.5	4	14	19	23.5	18
15	12.5	14	6	15	19	23	16
				16	22	23	4

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback
 Maximum possible score, 25.
 Change % is between pre- and post-test.

Appendix G – Likert questionnaire results**Error-corrected feedback group**

Student	1	2	3	4	5	6	7	8	9a	9b
1	2	2	2	2	2	2	18	8	Yes	5 weeks
2	2	2	2	1	1	2	21	9	Yes	1 month
3	2	1	2	2	-1	2	20	8	Yes	1 month
4	1	1	2	1	0	1	18	9	Yes	10 days
5	0	2	2	1	1	2	19	10	yes	1 week
6	1	1	1	2	1	1	20	10	n/a	n/a
7	1	1	1	1	-1	0	18	7	Yes	2 weeks
8	1	2	1	2	1	2	18	9	Yes	1 month
9	1	1	0	0	1	1	19	7	n/a	n/a
10	1	1	1	2	1	1	18	9	Yes	2 weeks
11	1	1	1	1	2	2	18	8	Yes	10 days
12	1	1	1	2	1	0	20	8	Yes	Can't remember
13	1	1	1	2	1	2	18	n/a	Yes	3 weeks
14	1	1	1	1	1	2	18	8	Yes	1 month
15	1	1	1	1	1	1	19	9	Yes	2 weeks

Dialogue journal writing group

Student	1	2	3	4	5	6	7	8	9a	9b
1	1	1	1	1	1	0	18	8	Yes	3.5 years
2	2	2	2	2	1	1	18	10	Yes	2 weeks
3	1	0	0	2	1	0	18	8	Yes	2 weeks
4	1	1	1	0	1	1	18	7	Yes	10 months
5	0	1	1	2	2	-1	19	6	Yes	1 year
6	1	1	1	1	1	1	19	12	Yes	4 years
7	0	1	1	2	-1	1	18	6	No	n/a
8	0	2	1	2	2	2	19	6	Yes	4.5 years
9	1	2	1	2	1	2	18	8	Yes	2 weeks
10	1	1	1	2	1	1	18	12.5	Yes	3 years
11	2	2	2	2	1	1	20	13	No	n/a
12	-1	0	1	2	2	0	20	10.5	Yes	4 years
13	1	1	1	2	1	1	20	12	Yes	5.5 years
14	2	2	1	2	1	1	18	13	Yes	3 weeks
15	2	2	2	2	2	2	18	10.5	Yes	6 years
16	2	1	2	2	2	1	23	11	No	n/a

Appendix H – Analytic scores for language in use (vocabulary choice and usage)

ECF group			DJW group		
Student	Pre-test	Post-test	Student	Pre-test	Post-test
1	3	3	1	4	4.5
2	2.5	3.5	2	3	4
3	2	3	3	3.5	4.5
4	3	3	4	3	4
5	2.5	3	5	3.5	4
6	3	3	6	4	4
7	3	3	7	3	4.5
8	3	4	8	4	5
9	3	3.5	9	3	4.5
10	2.5	2	10	4.5	4.5
11	2	3	11	4	4.5
12	2.5	3	12	3.5	5
13	3	4	13	3	4
14	3	3	14	4	5
15	2	3	15	3.5	4
			16	5	5

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback
 Maximum possible score, 5.

Appendix J – Total number of words produced

ECF group			DJW group		
Student	Pre-test	Post-test	Student	Pre-test	Post-test
1	270	251	1	299	257
2	136	321	2	258	300
3	71	244	3	156	194
4	185	245	4	195	297
5	83	195	5	216	290
6	173	267	6	240	279
7	145	194	7	235	256
8	138	387	8	337	313
9	130	285	9	221	237
10	191	166	10	409	286
11	124	193	11	270	282
12	154	257	12	325	338
13	193	314	13	303	196
14	206	258	14	249	364
15	116	205	15	258	322
			16	115	223

Abbreviations: DJW, daily journal writing; ECF, error-corrected feedback