

The Effects of Unfocused Error-Code WCF and Focused Metalinguistic Explanation on EFL Writing Accuracy

Yoshimasa Ogawa

Abstract

The present study compares the effects of (a) unfocused error-code corrective feedback alone and (b) a combination of unfocused error-code feedback and focused metalinguistic feedback on Japanese EFL students' accuracy in writing. Participants were three groups of English majors at Japanese universities, and they engaged in multi-draft writing tasks over a semester or an academic year. The changes in the number of grammatical errors as a function of WCF (written corrective feedback) were measured. The results showed that unfocused error-code WCF did not have any positive effect on the participants' language acquisition when it was administered singly. Comprehensive feedback kept participants from attending closely to individual grammatical forms and thus canceled out the positive features of error-code feedback. The combined form of feedback had a positive effect but not to a conclusive degree. The pedagogical implication was that learners must be guided to attend to feedback autonomously, and opportunities for new writings must be further increased.

Introduction

The influence of written corrective feedback (WCF) on L2 writing accuracy has been a major research issue in applied linguistics. For EFL or ESL teachers, checking, and providing feedback on, a large number of drafts is a daunting job, and it requires them of long experience and high-level expertise to offer appropriate written feedback in an efficient and effective way. As symbolized by Truscott's (1996, 1999) famous proposal to completely dispense with form-focused feedback, some L2 teachers may question the wisdom of investing time and effort for the provision of feedback, to which their students may not pay close attention. However, many studies including Ferris (1999, 2006) and Bitchener (2008) have produced evidence that form-focused feedback is likely to facilitate L2 learners' acquisition of grammatical forms in writing activities. Thus, the crucial issue is how to determine the best form of WCF—or the best combination of different types of WCF—for each target learner group in each particular learning context.

One major criterion for selecting the appropriate type of WCF is whether teacher feedback should be direct or indirect (Lalande, 1982; Chandler, 2003; Bitchener, Young, & Cameron, 2005; Sheen, 2007; Bitchener, 2008; Van Beuningen, De Jong, & Kuiken, 2012). Another important point to consider is whether WCF should be focused on one particular grammatical form or provided comprehensively on all grammatical errors that learners make (Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, Wright, & Moldawa, 2009). At the same time, it is important to remember that a combination of several different approaches is likely to be more

effective than the use of a single feedback technique (Bitchener et al., 2005; Bitchener & Knoch, 2010) because the multiple modes of feedback can compensate for the weaknesses of each other. The present study strove to evaluate the effectiveness of one multi-faceted WCF plan designed for first-year English majors enrolled in EFL writing courses at Japanese universities.

Earlier research findings have tended to indicate that indirect WCF is more useful and effective than direct correction in that the former engages L2 learners in cognitive processing. Two major types of indirect feedback are: (a) use of an error coding system (Lalande, 1982; Ferris, 2006; Van Beuningen et al., 2012) and (b) provision of metalinguistic explanation (Sheen, 2007; Bitchener & Knoch, 2009, 2010; Shintani & Ellis, 2013). Orthographic enhancement (e.g., circling or underlining ungrammatical parts) is yet another technique commonly utilized in everyday classroom situations, but it is more interesting to thoroughly evaluate the effectiveness of error codes and metalinguistic explanation which are more elaborate. The present study analyzed the functions of (a) comprehensive indirect error-code WCF alone and (b) a combination of unfocused error-code WCF and metalinguistic WCF focused on a few selected grammatical forms.

Regarding the role of focused WCF, the past studies had been mostly cross-sectional and, thus, focused on only one or two grammatical forms such as the English article system or simple past tense (Bitchener et al., 2005; Bitchener & Knoch, 2009). However, in real-world writing, L2 learners would be required to use all grammatical errors correctly (Ferris, 2011), and EFL writers often indicate their strong desire for feedback on most grammatical items (Ogawa, 2017). Consequently, in this longitudinal classroom-research study, the focused metalinguistic explanation was administered on several grammatical forms. That is, although the teacher provided metalinguistic explanation on one form at each session, different class sessions covered different error types. As mentioned above, error-code WCF was provided on all grammatical errors.

The research purposes were: (a) to evaluate the effects of unfocused, indirect WCF using an error coding system on Japanese EFL students' acquisition of grammatical forms and (b) to determine the degree to which focused metalinguistic WCF might enhance the effects of error-code feedback. The results of the study, when combined with the findings from other research studies in different EFL or ESL contexts, were believed to contribute to the overall development of more efficient and practical WCF approaches.

Literature Review

This section reviews some of the earlier studies on the effects of either error-code WCF or metalinguistic explanation on writing accuracy and those related to focused or unfocused feedback.

Error Code WCF and Metalinguistic WCF

Many researchers have provided evidence for the positive effects of error-code feedback on L2 learners' writing accuracy. For example, Lalande's (1982) study, involving a group of intermediate German-as-a-Foreign-Language students at a US university ($N=60$), showed that the group that

received indirect WCF using an error coding system gained greater accuracy than the group that received direct error correction in short-essay writing. Ferris (2006) evaluated error-code WCF's long-term influence on ESL students' essay writing in her study with 92 ESL students at a US university. As teacher, she indicated the location and type of each error that her students made, using codes for 15 error types. The participants made significantly fewer errors in their fourth essays than in the first essays, evidencing a positive long-term effect. Lalande and Ferris supported the position that error-code feedback was more effective than direct error correction.

Ferris and Roberts (2001) compared the effects of error-code feedback (explicit indirect feedback) and underlining (simpler indirect feedback). The types of errors covered were: verb errors, noun-ending errors, article errors, wrong word choice, and sentence-structure errors. The results showed that explicit indirect feedback was not necessarily more effective than the simpler feedback, but the group that received coded feedback outperformed the control group that did not receive any form-focused feedback. Furthermore, the results of a questionnaire survey indicated that the most popular error correction technique was for the teacher to mark errors and label them with codes.

Likewise, Van Beuningen et al. (2012) investigated the effects of error-code WCF and direct error correction on L2 Dutch learners' abilities to write accurately on a new topic. Participants, 268 secondary school students learning Dutch as their second language, wrote a short description of the biological features of an insect. The results showed that WCF in general (including direct correction and coded feedback) helped the participants gain greater accuracy in new pieces of writing, although the study did not demonstrate the superiority of coded feedback over direct error correction.

Overall, the studies above have shown that WCF using a coding system played a positive role for L2 learners' accuracy in writing, whereas there has been a dispute over the optimal explicitness of indirect feedback to be provided.

On the other hand, the usefulness of metalinguistic explanation has also been demonstrated by many L2-writing research studies. Sheen (2007) compared the effects of (a) direct correction only and (b) direct correction plus metalinguistic feedback on the acquisition of the English articles. Participants, 91 ESL students at a US language program, were divided into two experimental groups (receiving either direct correction alone or direct correction plus metalinguistic explanation) and a control group (receiving no feedback). The comparison between the three groups' means at pretest, immediate posttest, and delayed posttest indicated that the metalinguistic feedback contributed to learners' long-term acquisition of the target form.

Bitchener and Knoch (2009) recruited 52 low-intermediate ESL students in New Zealand as participants and evaluated the effects of metalinguistic explanation on their acquisition of the English definite and indefinite articles. They compared the effectiveness of three treatments: (a) written meta-linguistic explanation and an oral form-focused review of it, (b) written meta-linguistic explanation alone, and (c) error circling. All of these three experimental groups

outperformed the control group that received no feedback, although there was no significant difference between the three experimental groups. Then, in their follow-up study, Bitchener and Knoch (2010) evaluated the effectiveness of the same three types of WCF again and reported that the groups that received written metalinguistic explanation with or without an oral form-focused review achieved significantly greater accuracy than error circling.

Shintani and Ellis (2013) compared the effects of (a) direct error correction and (b) metalinguistic explanation on ESL learners' accurate use of the English indefinite article and demonstrated that metalinguistic explanation was more effective than direct feedback. Participants, a group of low-intermediate ESL students ($N=49$) in the United States, were divided into two experimental groups (receiving either metalinguistic feedback or direct correction) and one control group (receiving no feedback) and were engaged in a written picture-description task. The two experimental groups subsequently produced two new pieces of writing. The results showed that only the metalinguistic group outperformed the control group in accuracy, although their accurate use was not maintained until the delayed writing test two weeks later.

Focused and Unfocused WCF

As mentioned in the Introduction, most of the studies that assessed the influence of selective WCF on L2 learners' language acquisition were focused on a very small number of grammatical forms. Sheen (2007), Bitchener (2008), Bitchener and Knoch (2008), Bitchener and Knoch (2009), Bitchener and Knoch (2010) investigated the effects of WCF on ESL students' acquisition of the English article system alone. The results of these studies showed that experimental groups receiving various types of WCF (oral meta-linguistic explanation, written meta-linguistic explanation, direct error correction) consistently outperformed the control group that received no feedback in the acquisition of this target form. Bitchener et al. (2005) evaluated the effects of (a) explicit feedback only and (b) explicit written feedback and teacher-student conference on three grammatical forms (the definite article, the simple past tense, and prepositions) and found that the provided feedback had greater impact on the rule-governed features than on prepositions.

On the other hand, studies that contrasted the effects of focused and unfocused WCF have been scarce. Ellis et al. (2008) engaged a small group of Japanese EFL students in narrative writing tasks and measured the effects of focused and unfocused corrective feedback on their accurate use of grammatical rules in English. The participants were divided into three subgroups. One experimental group received WCF only on the definite and indefinite articles ($n=11$), and the other experimental group received WCF not only on the articles but also on other grammatical features ($n=13$). The control group received no feedback ($n=11$). Although both focused WCF and unfocused WCF groups made a significant improvement from pretest to posttest—both on an error correction test and a narrative-writing test—and also outperformed the control group, the study did not show whether focused feedback was more effective than unfocused feedback or vice versa.

Sheen et al. (2009) recruited 80 ESL students at a US college in the Washington DC area and compared the performance by (a) Focused Group that received direct WCF on the English articles alone, (b) Unfocused Group that received feedback on the articles as well as other grammatical structures, including copular *be*, regular past tense, irregular past tense, and prepositions, (c) Writing Practice Group that engaged in writing practice but received no feedback, and (d) Control Group that only took the tests. The ANOVA results showed that FG had a significantly greater gain than UG and CG between pretest and immediate posttest. At delayed posttest, FG performed better than CG, whereas WPG also outperformed CG. The implications were that focused WCF can contribute to grammatical accuracy in L2 writing but unfocused WCF is not pedagogically more useful than writing practice itself. Furthermore, FG gained greater accuracy than UG in the use of not only the articles but also the other grammatical structures. Sheen et al.'s interpretation was that focused WCF would urge learners to pay attention to non-target grammatical forms as well and monitor their use of all structures.

Overall, whereas the past studies have provided strong evidence for the positive effects of error-code WCF and metalinguistic WCF, the findings on the effectiveness of focused and unfocused WCF have not been conclusive. Furthermore, the interrelationship between these two factors has never been fully explored.

The following two research questions were put forward.

Research Question 1: Does unfocused error-code WCF facilitate Japanese students' acquisition of grammatical forms in the context of university EFL courses?

Research Question 2: To what extent does focused metalinguistic WCF enhance the effects of the error-code WCF on the acquisition of grammatical forms?

There was no a priori hypothesis for either research question.

Method

Site and Participants

This study was conducted as an action research study (Wallace, 1998; Burns, 2005) in order to improve the quality of the EFL courses that the researcher himself taught at two Japanese universities (hereafter referred to as Universities A and B, or simply UA and UB). A pseudo-experimental design was used, involving two intact classes at University A (University A1 and A2; or simply UA1 and UA2) and one intact class at University B. The participants at both universities were first-year English majors enrolled in basic paragraph- and essay-writing courses. The majority had not received any systematic paragraph-writing training in high school, and, therefore, the influence of prior paragraph-writing experience was mostly, if not entirely, controlled for. At the same time, it is also noteworthy that students in these EFL groups, who had received intensive grammar-focused instruction in junior high school and high

school, were familiar with grammatical rules and terms and were able to respond to the provision of error codes or metalinguistic explanations.

University A1 included 21 female first-year students enrolled in a reading-and-writing course in the spring semester of 2016. University A2 included 18 female first-year students enrolled in the reading-and-writing course in the fall of 2016. The fall course was a continuation of the spring course, but the classes were regrouped based on a placement test administered at the end of the spring semester; teachers, including the author of this paper himself, were placed in charge of different groups in the spring and the fall. In either semester, the class met twice a week, and the course ran for four months, providing 30 90-minute class sessions. The group at University B included 22 students (13 men and nine women) enrolled in a one-year EFL writing course in 2016. The class met once a week and had a total of 30 90-minute sessions.

Instructional Treatment

The pedagogical structures of the courses at the two universities were very similar. The students learned the basic principles, techniques, expressions, and conventions to write paragraphs and short essays, although the two schools' policies required different textbooks and entailed certain class administration constraints.

The course at University A was designed to improve learners' reading comprehension and paragraph-writing skills at the same time. The textbook assigned was *Unlock: Reading and Writing Skills 3* (Westbrook, 2014). The course covered Units 1 through 5 in the spring and Units 6 through 10 in the fall. The rhetorical patterns covered in the spring were: comparison and contrast, description, personal opinion, problem solution, and cause and effect. Those covered during the fall were: problem solution, advantage and disadvantage, personal opinion, explanation, and process. Class sessions were held in a regular classroom, and instruction was conducted mostly in English, providing the participants with opportunities for oral/aural English practice.

The textbook used at University B was *Get Your Message Across: Writing Communicative Paragraphs* (Jimbo, Elwood, Morita, Watanabe, Yamada, & Yoffe, 2008), which was prepared for paragraph- and essay-writing training in English. There were no exercises or questions intended primarily for reading comprehension practice. However, a set of sample writings, along with questions and simple exercises for understanding paragraph structures, were provided for each target rhetorical pattern. The rhetorical patterns covered during this year-long course were: time order, space order, process and direction, cause and effect, exemplification, definition, classification, and comparison and contrast. Class sessions were held in a computer-equipped room. As at University A, classroom instruction was conducted mostly in English.

The teacher assigned a total of five writing tasks per semester at University A, and the participants submitted three typed drafts for each assignment. At University B, the teacher assigned eight major writing assignments during the year (i.e., four assignments per semester) and one additional writing task as a final writing test. The participants submitted three typed

drafts for each of the eight major assignments and one draft for the final paragraph writing.

Written Feedback

Unfocused error-code WCF. The teacher provided indirect form-focused feedback on the participants' first and second drafts for each writing assignment both in the spring and the fall at both universities. The error types were indicated using codes (e.g., *VT* for verb-tense problems and *SV Agr* for subject-verb agreement problems). The teacher also offered content-based feedback whenever he found it necessary. Then, on the third draft, he corrected the remaining grammatical errors directly and gave grades based on the quality of the final draft and the efforts that the students had made during the drafting process. The participants were encouraged to try to revise their entire texts at each draft by experimenting with different expressions and sentence structures, adding new ideas, and/or restructuring the entire paragraph, instead of simply repairing the grammatical errors indicated by error codes.

Form-focused corrective feedback was provided on every draft based on the assumption that it was more effective to provide a certain amount of form-focused feedback continually (Fathman & Whalley, 1990; Ferris, 1995; Ashwell, 2000; Hartshorn, Evans, Merrill, Sudweeks, Strong-Krause, & Anderson, 2010). The students who took risks to improve their texts drastically tended to make additional errors on new drafts, and the teacher needed to provide follow-up form-focused feedback. Furthermore, for the students who made many grammatical errors, it was sometimes more appropriate to indicate some of the errors on the first draft and the others on the second draft so that they would not be overwhelmed.

Table 1 displays the types of errors on which the teacher provided form-focused feedback. The codes that the teacher used are given in parentheses. The error types can be categorized as

Table 1

The Types of Errors on Which WCF Were Provided

Treatable Errors

word order (WO), subject-verb agreement (SV Agr), pronoun agreement (Pro Agr), verb tense (VT), verb form (VF), noun form (NF), word form (WF), singular-plural problem (#), run-on sentence (run-on), fragment (Frag), voice (Vo), mechanics (including punctuation, capitalization, etc.; Mec), spelling (Sp), paragraphing error (Para),

Untreatable Errors

wrong word choice (WW), word missing (WM), unnecessary word (UnW), sentence structure (S/Str), idiomatic structure (Id), articles (Art)

Stylistic Errors

informal usage or expressions (Inf), sentence-initial conjunctive (In-Conj), redundant (Red), ambiguous (Amb), awkward (Awk)

Note. The underlined were the target forms for metalinguistic explanation.

treatable and untreatable errors (Ferris, 2011), as well as stylistically imperfect forms. However, whereas Ferris categorized the article system as treatable, this study treated it as untreatable error because it involved a large number of sub-rules and was extremely difficult for Japanese EFL students to master.

Error-code WCR plus metalinguistic explanation. In the fall, the teacher also provided University A2 and University B with metalinguistic explanation for a few selected grammatical forms, holding a mini-lecture. In other words, some of the frequent error types were highlighted during class sessions, in addition to the provision of error-code feedback to individual writers. He presented a set of example sentences that contained errors related to the target grammatical form and guided the participants to find, and correct, the errors corroboratively in small groups. Subsequently, he randomly called on a few students to indicate what the problematic parts were and how the errors could be corrected and, in the end, described the target grammatical rule explicitly. This mini-lecture took about 10–15 minutes per grammatical item. These consciousness-raising activities involving metalinguistic feedback were held for one selected grammatical form at a time; a total of seven grammatical forms were selected for such treatment.

The grammatical error types on which metalinguistic explanation was provided included: (a) noun form (i.e., a countable noun must be either preceded by an indefinite article or have the plural suffix on its end), (b) subject-verb number agreement, (c) definite and indefinite articles (*the* or *a/an*), (d) sentence fragment, (e) run-on sentence, (f) inappropriate use of conjunctives in sentence-initial position, and (g) active and passive voice. However, the number of errors that the participants made on the voice was very small, and, thus, this item was excluded from statistical analysis.

Analysis

A pretest-posttest design was used to measure the participants' accuracy in writing. Because of class administration constraints, it was not possible to conduct a delayed posttest. The positive effect of WCF on the participants' use of correct grammatical forms in writing was operationalized as the decrease, or increase, in the number of errors they made after receiving feedback over the semester. For each semester, the first draft for the first paragraph-writing task served as a preliminary writing test (hereafter referred to as *pretest*). The first draft for the final writing task—or the final writing test—served as a post writing test (hereafter as *posttest*). The difference between the numbers of errors the participants made at pretest and at posttest was believed to reflect their improvement, or regression, as a function of teacher written feedback.

The length of their paragraphs differed from person to person and from paragraph to paragraph. Consequently, the numbers of errors per 100 words were computed, and these adjusted numbers of errors were used as scores for statistical analyses. Then, a two-way mixed ANOVA was conducted to evaluate the effects of error-code WCF. The within-subjects factor was *test* with two levels (pretest and posttest), and the between-subjects factor was *school* with two levels (University A1 and University B). The dependent variables were the two groups' error

scores at pretest and posttest.

In order to understand the extent to which the error-code WCF influenced the participants' accurate use of six individual grammatical forms, the numbers of errors that the entire group in the spring (i.e., UA1 combined with UB) made on each of the selected grammatical forms at pretest and posttest were compared by performing six *t*-tests.

The effects of the two different types of WCF (i.e., (a) unfocused error-code WCF alone and (b) combination of unfocused error-code WCF and focused metalinguistic WCF) on writing accuracy was operationalized as the decrease, or increase, in the numbers of errors that the participants made after receiving either of the two types of feedback. Two different two-way ANOVAs were performed.

The first ANOVA (a mixed ANOVA) compared University A1's (error code WCF) and University A2's (error code plus metalinguistic WCF) gains, or losses, in the error scores. The within-subjects factor was *test* with two levels (pretest and posttest). The between-subjects factor was *treatment* (or *group*) with two levels (error code alone and error code plus metalinguistic). The dependent variables were each group's error scores at pretest and posttest. To assess the degree to which the two groups (UA1 and UA2) improved from pretest to posttest on each of the six grammatical forms, six *t*-tests were conducted to compare the difference means between the two groups.

Regarding University B, a two-way repeated-measures ANOVA was performed with *test* as the first within-subjects factor with two levels (pretest and posttest) and *treatment* (or *semester*) as the second within-subjects factor with two levels (error code alone and error code plus metalinguistic explanation). The dependent variables were the participants' error scores at each test. To evaluate the extent to which UB gained accuracy on individual grammatical forms after receiving either type of WCF, six *t*-tests were conducted, comparing the difference means for the spring and those for the fall.

The level of significance was set at $\alpha=0.05$ for all statistical analyses. Holm's sequential Bonferroni adjustment was used to prevent Type I errors in all *t*-tests.

Results

Evaluation of Error-Code WCF

In order to evaluate the effects of error-code WCF on EFL learners' writing accuracy, UA1 students' and UB students' error means in the spring were compared. Table 2 displays the descriptive statistics for the participants' performance on all the grammatical items on which they received WCF. Originally, UA1 had included 21 students, but it was reduced to 20 because one student, whose *z*-score was above ± 3.29 , was excluded from analyses (Field, 2009). UB included 22 students.

The mean for UA1 lowered from 0.41 ($SD=0.17$) at pretest to 0.35 ($SD=0.17$) at posttest, suggesting a minor but positive effect of error-code WCF on the participants' accuracy on all grammatical forms. On the other hand, the mean for UB increased from 0.23 ($SD=0.09$) to 0.26 ($SD=0.13$), indicating a slightly negative effect. It must be noticed, however, that UB's error

Table 2*Descriptive Statistics for UA1 vs UB Error Scores in Spring (All Grammatical Forms)*

			Pretest	Posttest
UA1	<i>M</i>		0.41	0.35
	95% CI	Lower Bound	0.32	0.27
		Upper Bound	0.48	0.43
	<i>SD</i>		0.17	0.17
	Skewness		1.04	0.93
	<i>SES</i>		0.51	0.51
	Kurtosis		0.56	1.26
	<i>SEK</i>		0.99	0.99
UB	<i>M</i>		0.23	0.26
	95% CI	Lower Bound	0.19	0.21
		Upper Bound	0.27	0.32
	<i>SD</i>		0.09	0.13
	Skewness		0.25	0.42
	<i>SES</i>		0.49	0.49
	Kurtosis		-0.73	-0.34
	<i>SEK</i>		0.95	0.95

Note. UA1, $n=20$; UB, $n=22$.**Table 3***Univariate Test Results of Two-Way ANOVA for UA1 and UB in the Spring (All Grammatical Forms)*

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
Test	1	0.002	0.002	1.52	0.70	0.01
Test x School	1	0.04	0.04	2.93	0.10	0.07
Residual	40	0.48	0.01			
School	1	0.36	0.36	12.02	0.01	0.23
Residual	40					

Note. $\alpha=0.05$.

scores were noticeably lower than UA1's at the beginning of the semester.

Table 3 presents the mixed ANOVA results for the mean-score comparison between UA1 and UB. Neither the *test* main effect nor the *test x school* interaction effect was statistically significant. However, the *school* main effect was significant, $F(1, 40)=12.02$, $p=0.01$, $\eta^2=0.23$. The effect size was large, indicating that the factor accounted for 23 % of the variance.

UB outperformed UA1 both at pretest and posttest. Then, the pretest-posttest mean differences were compared between the two schools. UA1's change in the error mean from pretest to posttest was -0.06 ($SD=0.52$), whereas UB's error mean rose by 0.03 ($SD=0.12$). The *t*-test result showed that there was no statistically significant difference, $t(40)=1.71$, $p=0.09$; the effect size was large, $d=5.15$.

Table 4 displays the descriptive statistics for the participants' performance on the six grammatical forms, which were selected for the purpose of between-treatments analysis in the fall. UA1 had originally included 21 students, but it was reduced to 19 because two students, whose z -scores were above ± 3.29 , were excluded from analysis. UB had first included 22 students, but it was reduced to 21 after the same adjustment.

UA1's error mean increased from 0.30 ($SD=0.31$) at pretest to 0.40 ($SD=0.35$), and UB's error mean also increased from 0.15 ($SD=0.16$) to 0.27 ($SD=0.18$). UB's means were lower than UA1's at both tests, and UA1 showed more individual differences.

Table 4

Descriptive Statistics for UA1 vs UB Error Scores in Spring (Six Grammatical Forms)

			Pretest	Posttest
UA1	<i>M</i>		0.30	0.40
	95% CI	Lower Bound	0.15	0.23
		Upper Bound	0.45	0.56
	<i>SD</i>		0.31	0.35
	Skewness		0.76	0.63
	<i>SES</i>		0.52	0.52
	Kurtosis		-0.30	-0.87
	<i>SEK</i>		1.01	1.01
UB	<i>M</i>		0.15	0.27
	95% CI	Lower Bound	0.08	0.19
		Upper Bound	0.23	0.35
	<i>SD</i>		0.16	0.18
	Skewness		0.81	0.23
	<i>SES</i>		0.97	0.50
	Kurtosis		-0.40	-0.90
	<i>SEK</i>		0.97	0.97

Note. UA1, $n=19$; UB, $n=21$.

Table 5

Univariate Test Results of Two-Way ANOVA For UA1 and UB in the Spring (Six Grammatical Forms)

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
Test	1	0.22	0.22	4.07	0.051	0.10
Test x School	1	0.03	0.03	0.53	0.82	0.001
Residual	38	2.10	0.06			
School	1	6.23	6.23	4.81	0.035	0.11
Residual	38					

Note. $\alpha=0.05$.

The ANOVA results shown in Table 5 indicated that there was a significant *school* main effect, $F(1, 38)=4.81$, $p=0.035$, $\eta^2=0.11$. UA1's pretest-posttest mean was 0.35 ($SD=0.05$), and

UB's pretest-posttest mean was 0.21 ($SD=0.4$); overall, UB made fewer errors than UA1.

The *test* main effect was very close to the significant level, $F(1, 38)=4.07$, $p=0.051$, $\eta^2=0.10$. The pretest mean for the two schools was 0.22 ($SD=0.25$), and the posttest mean was 0.33 ($SD=0.28$). That is, the unfocused error-code WCF had a nearly significant negative effect on the participants' overall accuracy on the six grammatical forms. The *test* x *school* interaction effect was not statistically significant.

Subsequently, six *t*-tests were performed in order to understand the error-code WCF's effect on individual grammatical forms. As shown in Tables 6 and 7, error-code WCF had a significantly negative effect on the participants' accuracy on *subject-verb agreement*, $t(39)=-3.63$, $p=0.001$ after Holm's sequential Bonferroni adjustment ($\alpha=0.008$); it had a nearly significant negative effect on *noun form*, $t(39)=-2.32$, $p=0.026$ ($\alpha=0.01$). These are two of the major grammatical forms on which participants tended to make frequent errors, and the *t*-test results suggested that they made more errors after receiving unfocused error-code WCF on these items.

Table 6

UA1 and UB Means at Pretest and Posttest (Six Grammatical Forms)

	Test	<i>M</i>	<i>SD</i>
SV Agreement	Pretest	0.30	0.19
	Posttest	0.49	0.76
Noun Form	Pretest	0.10	0.31
	Posttest	0.42	0.73
Article	Pretest	0.95	1.15
	Posttest	0.72	0.80
Run-on Sentence	Pretest	0.05	0.31
	Posttest	0.04	0.17
Fragment	Pretest	0.10	0.30
	Posttest	0.07	0.26
Sentence-Initial Conjunctive	Pretest	0.11	0.48
	Posttest	0.24	0.44

Note. UA1, $n=19$; UB, $n=21$.

Table 7

t-Test Results for Pretest-Posttest Differences on Six Grammatical Forms

	<i>t</i>	<i>p</i>
SV Agreement	-3.63	0.001
Noun Form	-2.32	0.026
Article	1.27	0.21
Run-on Sentence	0.19	0.85
Fragment	0.44	0.66
Sentence-Initial Conjunctive	-1.31	0.20

Note. $df=39$.

Evaluation of the Two WCF Treatments

First, the means for UA1 and UA2 were computed to compare the effects of the error-code WCF and the error-code-plus-metalinguistic WCF (Table 8). A mixed ANOVA was performed to determine the level of significance.

Table 8
Descriptive Statistics for UA1 vs UA2 Comparison (Six Grammatical Forms)

			Pretest	Posttest
UA1	<i>M</i>		0.24	0.49
	95% CI	Lower Bound	0.10	0.20
		Upper Bound	0.38	0.78
	<i>SD</i>		0.28	0.59
	Skewness		1.26	2.40
	<i>SES</i>		0.54	0.54
	Kurtosis		1.51	0.30
	<i>SEK</i>		1.04	1.04
UA2	<i>M</i>		0.31	0.30
	95% CI	Lower Bound	0.18	0.16
		Upper Bound	0.44	0.44
	<i>SD</i>		0.25	0.26
	Skewness		1.08	1.26
	<i>SES</i>		0.56	0.56
	Kurtosis		1.93	1.25
	<i>SEK</i>		1.09	1.09

Note. UA1, $n=18$; UA2, $n=16$.

The n -size for UA1 was reduced from 21 to 18 because three participants, whose z -scores were above ± 3.29 , were deleted. Through the same procedure, the n -size for UA2 was reduced from 18 to 16. (The statistics for UA1 in Table 8 are slightly different from those in Table 4 because different students were eliminated as outliers.)

The error-score mean for UA1 increased from 0.24 ($SD=0.28$) at pretest to 0.49 ($SD=0.59$) whereas the mean for UA2 decreased slightly from 0.31 ($SD=0.25$) at pretest to 0.30 ($SD=0.26$) at posttest, implying that the combined form of WCF might have been somewhat more effective.

The univariate test results in Table 9 showed that the *test* main effect, the *group* main effect, and the *test* x *group* interaction effect were all not significant. However, there was a possibility that the treatments might have had positive effects on some individual grammatical forms. Consequently, UA1's and UA2's pretest-posttest mean differences on each of the six grammatical items were computed and compared (Table 10).

Table 9*Univariate Test Results of Two-Way ANOVA for UA1 and UA2*

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
Test	1	0.25	0.25	1.79	0.19	0.05
Test x Group	1	0.30	0.30	2.14	0.15	0.06
Residual	32	4.45	0.14			
Group	1	0.06	0.06	0.40	0.53	0.01
Residual	32	4.73	0.15			

Note. $\alpha=0.05$.**Table 10***Pretest-Posttest Mean Differences for UA1 and UA2 (Six Grammatical Forms)*

	Group	<i>M</i>	<i>SD</i>
SV Agreement posttest_SV Agreement pretest	A1	0.65	1.17
	A2	0.34	1.00
Noun Form posttest_Noun Form pretest	A1	0.49	0.80
	A2	-0.49	0.93
Article posttest_Article pretest	A1	-0.09	1.88
	A2	0.58	1.19
Run-on posttest_Run-on pretest	A1	0.04	0.16
	A2	-0.02	0.08
Fragment posttest_Fragment pretest	A1	0.29	1.42
	A2	-0.14	0.26
Initial Conjunctive posttest_Initial Conjunctive pretest	A1	0.14	0.57
	A2	-0.34	0.49

Note. UA1, $n=18$; UA2, $n=16$.

Although the overall effects of the two treatments were not significantly different, the negative t -values in Table 11 show that the participants benefited more from the combined WCF on all grammatical forms except for *article*. The difference was statistically significant for *noun form*, $t = -3.30$, $p = 0.002$, after Holm's sequential Bonferroni adjustment ($\alpha = 0.008$), and it was almost significant for *sentence-initial conjunctives*, $t = -2.60$, $p = 0.014$ ($\alpha = 0.01$).

Table 11*t-Test Results for UA1's and UA2's Mean Differences on Six Forms*

	<i>t</i>	<i>p</i>
SV Agreement posttest_SV Agreement pretest	-0.84	0.41
Noun Form posttest_Noun Form pretest	-3.30	0.002
Article posttest_Article pretest	1.22	0.23
Run-on posttest_Run-on pretest	-1.32	0.20
Fragment posttest_Fragment pretest	-1.20	0.24
Initial Conjunctive posttest_Initial Conjunctive pretest	-2.60	0.014

Note. $df=32$.

Second, a two-way repeated-measures ANOVA was conducted to compare UB's performance in the spring (receiving error-code WCF) and in the fall (receiving combined WCF) (Table 12).

Table 12

Descriptive Statistics for UB (Spring) vs. UB (Fall) Comparison (Six Grammatical Forms)

			Pretest	Posttest
UB (spring)	<i>M</i>		0.17	0.31
	95% CI	Lower Bound	0.09	0.20
		Upper Bound	0.25	0.42
	<i>SD</i>		0.18	0.25
	Skewness		0.75	1.55
	<i>SES</i>		0.49	0.49
	Kurtosis		-0.67	3.85
	<i>SEK</i>		0.95	0.95
UB (fall)	<i>M</i>		0.34	0.22
	95% CI	Lower Bound	0.25	0.15
		Upper Bound	0.43	0.29
	<i>SD</i>		0.20	0.15
	Skewness		0.34	1.50
	<i>SES</i>		0.49	0.49
	Kurtosis		-0.54	3.03
	<i>SEK</i>		0.96	0.95

Note. $n=22$.

The kurtosis values for the spring and fall posttests are both above +1.96, and, therefore, the results must be interpreted cautiously. However, UB's error mean decreased from 0.34 ($SD=0.20$) to 0.22 ($SD=0.15$) in the fall whereas its mean increased from 0.17 ($SD=0.18$) to 0.31 ($SD=0.25$) in the spring, suggesting that the combined WCF had more positive influence on the students' accuracy in writing.

Table 13

Univariate Test Results of Two-Way Repeated-Measures ANOVA for UB in Spring and Fall

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
Semester	1	0.04	0.04	1.09	0.31	0.05
Residual	21	7.77	0.04			
Test	1	0.002	0.002	0.05	0.82	0.002
Residual	21	0.71	0.03			
Semester x Test	1	0.37	0.37	11.60	0.003	0.36
Residual	21	0.68	0.03			

Note. $\alpha=0.05$.

The interaction effect shown in Table 13 was statistically significant, $F(11.60)$, $p=0.003$, $\eta^2=0.36$. The follow-up *t*-test showed that the decrease in error means from pretest to posttest for

the fall semester (i.e., for combined WCF; $M = -0.12$; $SD = 0.28$) was significantly greater than that for the spring semester (i.e., error-code WCF alone; $M = 0.14$; $SD = 2.33$), $t(21) = -3.40$, $p = 0.003$, $d = 0.16$.

Subsequently, the mean differences between the pretest and the posttest were compared on each of the six individual grammatical forms (Table 14).

Table 14

Pretest and Posttest Mean Differences for UB in Spring and in Fall (Six Grammatical Forms)

	Group	<i>M</i>	<i>SD</i>
SV Agreement posttest_SV Agreement pretest	Spring	0.55	0.78
	Fall	0.09	0.73
Noun Form posttest_Noun Form pretest	Spring	0.05	0.70
	Fall	-0.71	1.31
Article posttest_Article pretest	Spring	0.04	1.05
	Fall	-0.29	0.79
Run-on posttest_Run-on pretest	Spring	0.00	0.00
	Fall	0.03	0.14
Fragment posttest_Fragment pretest	Spring	-0.08	0.33
	Fall	0.19	0.63
Initial Conjunctive posttest_Initial Conjunctive pretest	Spring	0.27	0.43
	Fall	-0.04	0.49

Note. $n = 22$.

The t -test results in Table 15 showed that, although the participants' error scores increased (i.e., their accuracy decreased) slightly on *run-on sentence* and *sentence fragment*, they benefited more from the combined WCF on the other grammatical forms. The combined WCF had noticeably stronger effects on *SV agreement*, *noun form*, and *sentence-initial conjunctive* although the p -values did not reach the level of significance after Holm's sequential Bonferroni adjustment.

Table 15

t-Test Results for Six Forms for UB in Spring and Fall (Six Individual Forms)

	<i>t</i>	<i>p</i>
SV Agreement post_SV Agreement pre	-2.14	0.04
Noun Form post_Noun Form pre	-2.27	0.03
Article post_Article pre	-1.08	0.29
Run-on post_Run-on pre	1.00	0.33
Fragment post_Fragment pre	1.81	0.09
Initial Conjunctive post_Initial Conjunctive pre	-2.47	0.02

Note. $df = 21$.

Discussion

Research Question 1 was: Does unfocused error-code WCF facilitate Japanese students' acquisition of grammatical forms in the context of university EFL courses? The results showed that unfocused error-code WCF did not help improve the participants' accuracy in writing. The changes in error rate at neither UA1 nor UB were statistically significant, but, regarding *subject-verb agreement* and *noun form*, two of the most frequently observed error types, the participants' error means increased to a nearly significant degree. This may well be interpreted as evidence that unfocused error-code WCF does not contribute to EFL writers' language acquisition or that the instructional process involved some factor that had a detrimental effect on its positive function.

Lalande (1982), Ferris (2006), and Van Beuningen et al. (2012) have demonstrated that error-code WCF is a useful instrument for language teaching. On the other hand, Sheen et al. (2009), as one of the few research teams that compared the effects of focused and unfocused WCF, indicated that unfocused WCF was not pedagogically more useful than writing practice itself. Consequently, the weaknesses of unfocused feedback might have canceled out the positive effects of error-code feedback. The participants who were all first-year students, although a certain proficiency difference was recognized between the two universities, made a considerably large number of grammatical errors on the first draft of each writing assignment, and they might have been unable to attend to each and every grammatical error type that the teacher indicated. For future instruction, it is imperative to make instructional plans to prevent this trade-off situation.

Another possible cause of their lack of improvement was that the number of writing tasks (i.e., new topics for writing) was limited to only four or five per semester due to the class administration constraints, and the writing training lasted for only one semester, which was not long enough for language acquisition to take place. This problem is related to the first problem described above: if there are many target grammatical forms to attend to and few opportunities to use those forms repeatedly, it is quite possible that learners' attention will become tenuous. Incidentally, one rule that most participants mastered after finishing one or two writing tasks was *paragraph structure* (i.e., each paragraph must include a topic sentence, supporting sentences, and a concluding sentence, and there should be no blank space between sentences within the same paragraph) because they must apply this rule in every writing task.

Yet another possible factor that resulted in the increased numbers of errors was the participants' psychological reaction to assignments. At the beginning of the year, these first-year students paid close attention to the teacher's directions and concentrated on the given tasks, but, as they became accustomed to the task patterns, they became somewhat careless in finishing the assigned writings. The teacher encouraged participants, time and again during the semester(s), to edit the entire texts at each stage of drafting in terms of both form and content,

instead of mechanically repairing the ungrammatical parts. However, he observed, during class sessions and in the process of collecting drafts, that not all participants were enthusiastic about process-oriented writing or willing to spend a large amount of time on drafting.

Research Question 2 was: To what extent does focused metalinguistic WCF enhance the effects of the error-code WCF on the acquisition of grammatical forms? Although not conclusive, the results of this study produced evidence that the addition of metalinguistic explanation helped learners acquire the target grammatical forms to some extent. The comparative analyses between UA1 (receiving error-code WCF alone) and UA2 (receiving combined WCF) showed that, overall, there was no significant difference between the two group's performance. However, the combined WCF led to UA2 participants' greater accuracy on five out of the six target grammatical forms with the notable exception of *article*, which is absent in the Japanese syntax and known to be a difficult grammatical form for Japanese EFL learners to master. Particularly, UA2's error scores decreased significantly on *noun form* and almost significantly on *sentence-initial conjunctive*. Likewise, the analysis of UB's performance in the spring (error code only) and in the fall (combined WCF) also revealed that the mini-lectures involving metalinguistic explanation played a positive role. Overall, their error scores were significantly lower in the fall when they received the combined WCF. Regarding individual grammatical forms, they performed better in the fall on all items except for *run-on sentence* and *fragment*. Particularly, the differences were noticeably large on *subject-verb agreement*, *noun form*, and *sentence-initial conjunctive*.

One interesting point is that *noun form* and *subject-verb agreement* are two of the major treatable error types whereas *sentence-initial conjunctive* is a stylistically imperfect form. This study showed that the combined WCF exerted positive influence on learners' acquisition of both types of errors. However, there is a possibility that EFL learners might react differently to different categories of grammatical forms, and future research studies might focus on this issue as well.

Overall, the teacher may continue to use this form of combined WCF as a useful instructional technique. Whereas unfocused WCF failed to exert positive influence, the focused metalinguistic explanation, which had been designed to enhance the effects of error-code WCF, compensated for the drawbacks of unfocused feedback to a certain degree. It is acknowledged, however, that there is clearly room for improvement concerning the way to administer metalinguistic explanation itself. Mini-lectures that provide metalinguistic explanation might be administered more frequently, additional measures might be taken to draw learners' attention more closely to target pedagogical points, and the syllabus for their overall writing training might also be readjusted.

Conclusion

The study demonstrated that unfocused WCF using an error-code system did not contribute

to the participants' accuracy in writing but that the combination of error-code feedback and metalinguistic explanation had a positive effect on their learning of grammatical forms. Unfortunately, the positive evidence for the combined WCF itself was not conclusive, either, implying that the feedback approach used in the present study needs to be modified in order to make it fully functional.

The absence of conclusive evidence, particularly for unfocused error-code WCF, might suggest that Truscott's (1996) proposal that WCF is a waste of time and effort both on the teacher's and students' sides and should be abandoned is valid. However, a more realistic and constructive interpretation is that the WCF treatment administered in the present project was not intensive enough to improve their accuracy in one or two semesters. Although participants submitted a new draft every week, they worked on only four or five different topics per semester, and the cognitive processing that they experienced might have been limited. Consequently, one alternative pedagogical plan is to increase the number of writing assignments even at the expense of the traditional three-draft policy, e.g., requiring learners to submit two drafts, instead of three, in order to make room for additional writing assignments. Another alternative is to assign journal writing homework or utilize an automated writing feedback system, which increases learners' opportunities for writing without requiring teacher feedback. The participants at University UB were guided to write a few essays utilizing an automated feedback system during the year, but the emphasis was always on the paragraphs to submit to the teacher. Thus, the total number of paragraphs or short essays they finished was rather small. After all, WCF is an instrument to support L2 learners' writing practice, and the opportunities for writing themselves must be increased in one way or another.

Another possible cause of the participants' failure to acquire greater accuracy in writing was that they did not always pay close attention to grammatical forms when they wrote. The participants' drafts showed that many of them were making the same errors repeatedly even after receiving error-code and/or metalinguistic WCF on those items. One remedial approach might be for the teacher to intensify the explicit feedback on the major target forms, for example, by holding mini-lectures on common error types several times during the semester, not just once. However, overly elaborate or explicit feedback does not guarantee positive results because it can deprive learners of chances to notice their own problems and strive to solve them. Another alternative is to simply underline the types of errors on which participants have received metalinguistic explanation in earlier mini-lectures, instead of giving explicit feedback on the same types of errors repeatedly, and direct them to look back to the earlier lessons. Learners must be guided to autonomously engage in error correction.

Overall, the multi-layered WCF treatment proposed in this study might be revised into a more intensive type, but, at the same time, the syllabus for the EFL writing training itself needs to be readjusted. Sufficient opportunities for new writings are imperative, and learners should be guided to develop their own strategy to correct their errors and utilize teacher

feedback more consistently. Future research WCF might also be focused on learners' reaction to formed-focused feedback as well as on the forms of corrective feedback per se. EFL learners' grammatical accuracy in L2 writing can be improved only when the teacher and learners work in tandem with each other for this purpose.

References

- 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(3), 227-257.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102-118.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12(3), 409-431.
- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal*, 63(3), 204-211.
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207-217.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(3), 191-205.
- Burns, A. (2005). Action research: An evolving paradigm? *Language Teaching*, 38(2), 57-74.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267-296.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353-371.
- 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, U.K.: Cambridge University Press.
- Ferris, D. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly*, 29(1), 33-53.
- Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1-11.
- Ferris, D. R. (2006). Does error feedback help student writers?: New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contents and issues* (pp. 81-104). New York: Cambridge University Press.
- Ferris, D. R. (2011). *Treatment of error: In second language student writing* (2nd ed.). An Arbor: The University of Michigan Press.
- Ferris, D. R., & 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.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.
- Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 44(1), 84-109.
- Jimbo, H., Elwood, K., Morita, A., Watanabe, Y., Yamada, S., & Yoffe, L. (2008). *Get your message across:*

Writing communicative paragraphs. Tokyo: Nan'un-do.

- Lalande, J. F., II. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, 66(2), 140-149.
- Ogawa, Y. (2017). Japanese EFL writers' perceptions of, and reactions to, written corrective feedback. *Gakuen*, 917, 1-23.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255-283.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Different effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556-569.
- Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 22(3), 286-306.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-369.
- Truscott, J. (1999). The case for "the case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1-41.
- Wallace, M. J. (1998). *Action research for language teachers*. Cambridge: Cambridge University Press.
- Westbrook, C. (2014). *Unlock: Reading and writing skills 3*. Cambridge, UK: Cambridge University Press.

(小川 喜正 英語コミュニケーション学科)