Investigating Second Language Development in Writing Using the Extremely Short Story Task

ESS における第二言語ライティングの発達研究

Hiroaki KATO
加藤洋昭

Abstract

This study investigated whether Japanese university students show second language (L2) writing development underlying these behaviors; whether teachers' corrective feedback affects linguistic accuracy, complexity and fluency of written output during five essay writing tasks. Frequent errors during the writing tasks were also analyzed. Participants were 18 intermediate L2 writers, who completed five essay writing tasks. The writing tasks required a rigid submission rule. Participants were required to write an essay with exactly fifty words. Participants' writing behaviors were analyzed from the perspective of second language writing development. However, the present study disconfirms predictions regarding syntactic complexity, linguistic accuracy and fluency. Implications for research, evaluations of essay writing and pedagogy are discussed.

Keywords: Extremely Short Story, Accuracy, Complexity, Fluency, Second Language Development in Writing

Introduction

English education in Japan has changed drastically over the past five years. The current course of study for senior high school was implemented in April 2013 and essentially stipulates that English should be taught in English. The Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) emphasizes that English should be used to teach English. The Japanese government also produced a proposal to cultivate global human resources in 2012 requiring that young Japanese people acquire linguistic and communication skills, self-direction and positivity, and understanding of other cultures and a sense of Japanese identity (Cabinet Office, 2012, p. 8). In spite of these policies, the TOEFL results remain relatively poor. The results of the TOEFL iBT in 2016 show that Japan is ranking 27th out of 31 Asian countries (Educational Testing Service, 2017). The Japanese examinees scored 19 out of 30 points on the writing section, which was the highest score of the four skills, but its score is still ranked 27th out of 31 Asian countries. From these perspectives, for practical reasons, the investigation of writing research in formal language teaching is required to improve Japanese learners of English. The purpose of this study is to develop Japanese learners of English through the writing of extremely short stories consisting of exactly 50 words.

1. Extremely Short Story

The idea of an extremely short story (ESS) was proposed by Hassall in 2006. The ESS is a story or essay written in only 50 words. The writers can choose any topic, genre, or style. Some writers choose fiction and others non-fiction. The title and author's' name should be excluded from the word count, and the only restriction is that the story or essay be completed in 50 words.

This writing practice will be beneficial for Japanese learners of English for two reasons (Honna, Kirkpatrick, & Takeshita, 2018, p. 30). First, this limitation on the

number of words is very important because it requires learners to manage their use of English through word choice. For example, when one word is short or over the limit, learners need to deal with the problem using their lexical and grammatical knowledge. The learners will acquire English through these manipulations and learn how to express themselves in English. Second, this fifty-word rule enables non-native speakers of English to compete with native speakers of English. The learners will make sure they have developed a remarkable command of English through the writing practices. Honna (2008: 166-167) says foreign students from the United States concede that many Japanese works are better than theirs. The Japanese Association for Asian English (JAFAE) has conducted seven competitions so far from 2007 to 2012, and many Japanese students have received prizes in the competitions (see http://www.jafae.org/essc/). However, no studies have ever attempted to examine the effects of corrective feedback and sequence writing practices. This is the aim of the present study.

2. An Overview of Second Language Writing Research

In this section, several empirical studies on writing issues are reviewed. First, students' attitudes towards error correction are reported, followed by a review of a study on the effects of corrective feedback of writing.

2.1 Research Evidence on Students' Attitudes Towards Error Correction

Lee (2005) investigated L2 secondary students' perceptions, beliefs and attitudes about error correction. A total of 320 students from eight secondary schools in Hong Kong answered a questionnaire. Twenty-seven of the students also took part in an individual follow-up interview. She found that most of the teachers adopted comprehensive error feedback. The majority of students also preferred to receive comprehensive feedback, and 76% of students also wished their teachers to correct all errors. The students valued error correction and believed it would lead to improv-

ing their proficiency. Although the students were not sure whether they were making progress, they preferred direct, explicit feedback instead of indirect feedback. Her research showed that both teachers and students have a preference for direct and explicit feedback.

Lyster, Saito, and Sato (2013) also reported a tendency similar to the one found by Lee (2005); there is a clear trend for learners to express a preference for receiving corrective feedback. In other words, learners do not desire to have their errors ignored.

2.2 Research Evidence on the Effect of Types of Feedback

Robb, Ross, and Shortreed (1986) investigated the effect of salient feedback on revision writing. The hypothesis was set as follows: more salient error feedback would have a significant effect on improving the students' overall writing quality (Robb et al. p. 85). The participants were 134 Japanese college freshmen. The study was conducted over the academic year, from mid-April until mid-January. Participants were divided into four groups: (a) correction group, all errors completely corrected by the instructor; (b) coded feedback group, errors marked in an abbreviated code system; (c) uncoded feedback group, errors marked over with a yellow text-marking pen, which differed from the coded feedback in that the instructor only indicated the place of errors; and (d) marginal feedback group, the least salient method group, where the total number of errors was written in the margins of the student's paper. Learners in all groups were required to write expository, narrative and descriptive essays. After they received the instructor's feedback, the composition data were analyzed in terms of accuracy, complexity and fluency. The study found that the overt correction did not lead to more accuracy. However, the fluency of all four groups increased. The researchers noted that the weekly writing assignments created a practice effect. On the complexity measure, there were no statistically significant differences between the groups. Robb et al. (1986) assume that even the direct correction was still too obscure to help students to deal with the previous

errors. From these results, it appears that less time-consuming methods of directing student attention to error may be sufficient. Teachers should not expect that direct comprehensive feedback, a time-consuming method, would result in more accurate composition. This study showed that as the course progressed, students in all of the groups produced more complex structures. The researchers concluded that the improvement was independent of type of feedback. However, this study did not include a non-feedback control group, limiting its conclusions.

Bitchener, Young, and Cameron (2005) investigated whether the different types of feedback improved accuracy in four new pieces of writing over a 12-week period. The participants were 53 post-intermediate learners. They were divided into three groups, the first receiving direct written corrective feedback and a 5-minute student-researcher conference after each piece of writing, while the second group received direct written corrective feedback only. The third group received no corrective feedback. Direct written corrective feedback occurs when the teacher indicates that an error has been made and provides a correction, thereby not leaving the student to diagnose and correct it. The three linguistic errors that occurred most frequently during the first writing task (prepositions, past simple tense and definite article) were chosen as the feedback points. The study found that the group receiving direct written corrective feedback with individual conferences had greater accuracy on new pieces of writing. While their performance of past simple tense and the definite article improved, prepositions did not. However, when the three error categories were considered as a single group, the type of feedback did not have a significant effect on accuracy. The data showed that overall accuracy is not a linear and upward pattern of improvement. From this result, Bitchener et al. suggested that the investigation of feedback should be examined longitudinally.

Sheen (2007) investigated the effect of different feedback on ESL learners' acquisition of articles. The participants were 111 intermediate-level university students who were divided into three groups. In the direct-only correction group, a traditional error correction strategy, learners' errors were identified and corrected by

the instructor. In the direct metalinguistic correction group, learners' errors were identified, corrected and included metalinguistic comments that explained the correct form. The last group was a control group. The targets were articles, both indefinite and definite, selected for two reasons. First, the participants did not learn articles explicitly during the semester. Second, students commonly make errors with articles, but the rule explanation is complicated. The participants took a pretest. After the treatment, they took a posttest and delayed posttest. Two narrative tasks were used to elicit articles. To measure learners' performance, three tests were administered; a speeded dictation test, a narrative writing test and an error correction test. In the speeded dictation test, the participants listened to 14 sentences. Each sentence was read at a normal speed, and participants were required to write down what they heard as fast and precisely as possible. In the writing test, the participants were asked to narrate a story based on four sequential pictures. In the error correction test, the participants were asked to correct a phrase which was underlined and contained an error. The participants also took a 14-item language analytic ability test. The participants were given a glossary of words and sentences of artificial language along with their English translations. They were asked to choose the correct translation from the four choices. Throughout the test, the participants were required to deduce the rule and apply it to the choices. The study found that there was a significant difference between the direct metalinguistic group and control group in the delayed posttest. Concerning the writing test, there was a significant difference among the three groups in the delayed posttest. There was an also significant difference between the post and delayed posttests. This study showed that direct correction with metalinguistic comments was superior to direct correction without metalinguistic comments. Sheen (2007, p. 275) argues that while both direct correct feedback with and without metalinguistic comments may promote awareness as noticing, direct corrective feedback with metalinguistic comments promotes awareness with understanding. The study also found that learners who had high aptitude benefited more from both types of corrective feedback. This may imply that learners who have a higher aptitude for

language analysis are more likely to increase awareness from the corrective feedback. This study was unique in that it investigated not only the effect of feedback but also language aptitude.

Van Beuningen, De Jong and Kuiken (2012) investigated the effect of direct and indirect corrective feedback on learners' accuracy. They posed two questions regarding the role of corrective feedback. The first was whether corrective feedback would improve the learners' accuracy as a revision tool, and the second was whether corrective feedback would support long-term accuracy development and be demonstrated in new pieces of writing. The participants were 268 Dutch secondary school students. They were divided into two groups, higher level and lower level, due to a presumption that learners' proficiency would indicate their level of metalinguistic awareness. The researchers expected that higher-level students would make more corrections from indirect corrective feedback than lower-level students. This study set out two experimental treatments (direct corrective feedback and indirect corrective feedback) and two control conditions (self-correction and additional writing practice). It took six weeks to complete. In the first week, the participants took a receptive vocabulary test, a questionnaire concerning their language background and the first writing task. The first writing task served as a baseline of participants' accuracy, structural and lexical complexity. In the second week, participants received either direct or indirect corrective feedback. The self-correction group revised their text based upon the first draft handed in the previous week without any alteration. In the third week, all participants were asked to produce a new text with a new topic. In the sixth week, three weeks after the post-test, the participants were again asked to produce a new text on a new topic. As was expected, in the first session the first week of the pretest, the lower students committed more errors than did the higher-level ones. Concerning the revising tool, both the corrective feedback group and self-correction group decreased their error rate. The practice group, however, committed more errors than on the pretest. There was no interaction effect between group condition and language proficiency. The researchers emphasized the effects of feedback

as follows: “both corrective feedback treatments turned out to be significantly more beneficial than either of the control conditions” (Van Beunigen et al. 2012, p. 24). In terms of short-term learning, effects were investigated in the third week, and long-term learning effects were investigated in the sixth week. Both corrective feedback groups outperformed the two control groups. This study also found that only the direct corrective feedback group reduced the number of mistakes in the third week of posttest and sixth week of delayed posttest. The participants’ writing practice did not lead to simplified writing. There were no significant differences between groups on the measures of structural complexity and lexical diversity in the new writing text. This study found that comprehensive corrective feedback enables learners to improve their accuracy through the revision. This study also found that unfocused corrective feedback learners made fewer errors on new pieces of writing than learners who did not receive corrective feedback. Of note, only direct correction promoted grammatical accuracy, as mentioned earlier. In contrast, indirect corrective feedback improved non-grammatical accuracy in the delayed posttest. This study demonstrated the effect of corrective feedback.

3. The Present Study

3.1 Research Questions

The present study examines the effects of corrective feedback on measures of learner production on five sessions of ESS writings over 15 weeks. This study also examines the most frequent types of errors made by Japanese university undergraduate students. Two research questions are addressed in this study.

3.1.1 Research Question 1

What types of errors are made frequently in ESS writings by Japanese learners of English?

3.1.2 Research Question 2

Do the Japanese learners of English writing show development on ESS writings in terms of accuracy, complexity and fluency? The hypothesis is that as the learners become more proficient, they will write more accurately, grammatically and lexically complex sentences, as well as more fluently.

3.2 Participants and Research Design

The participants were 18 Japanese university undergraduates, aged between 19 and 21, attending a private university. All participants had completed six years of junior and high school English education, and one year of an English program at the university. When the participants were freshmen, they took courses such as Speaking, Listening, and Reading and Writing. The participants were judged to be at an intermediate level by the university.

The participants were sophomores and required to write ESS as their assignments. They were allowed to access English-Japanese or Japanese-English dictionaries. No time limit was set. During the semester, the participants wrote five ESSs every four weeks. They were also asked to write comments on the work in Japanese, their mother tongue. Corrective feedback and comments on the topic were then given by the researcher.

3.3 Measures

Learner production was assessed for accuracy, complexity, and fluency. To measure accuracy, a general measure of error free T-units (EFT) was used. A T-unit is defined as “one main clause plus whatever subordinate clauses happen to be attracted or embedded within it” (Hunt, 1966, p. 735, cited in Robinson, 1995, p. 110). Syntactic complexity was assessed by a measure of sentence nodes per T-units (SPT) and type-token ratio (TTR). Fluency was assessed by a measure of words per error-free T-unit (W/EFT). These general measures were used because a number of studies have found them to have a moderate correlation with proficiency (Wolfe-Quintero,

Inagaki, & Kim, 1998).

4. Results

4.1.1 Research Question 1. Frequent Errors

The first addressed research question concerned the types of errors frequently made by the Japanese university students. The errors were categorized into the following 16 types, in alphabetical order: agreement, article, conjunction, fragment, lexical errors, null subject, plural, preposition, pronoun, relative clause, singular, tense, 3rd singular, voice, word order, and other. Examples of errors are cited in the appendices. Table 1 shows the frequency of errors made by the Japanese students of English on five ESS writings. Lexical errors were made the most frequently. They include addition of inappropriate words, insertion of intangible words, errors of parts of speech, and usage (intransitive verb or transitive verb). The errors of article consisted of missing definite or indefinite articles.

Table 1

Frequency of Errors

Lexical Errors	Article	Tense	Pronoun	Plural	Agreement	Preposition	Voice
93	65	46	36	35	25	23	16
Conjunction	Fragment	Word Order	3rd Singular s	Null Subject	Singular	Relative Clause	Others
13	13	7	5	5	4	2	2

Table 2 shows the detail of lexical errors. Most of these were due to inappropriate word selection, as seen in this example: *What kinds of drink do you favorite the best?* The word “favorite” is not a verb but an adjective. However, it is used as a verb in this sentence. Thus, this was counted as an inappropriate word, as the word “like” should have been used instead.

Table 2

Details of Lexical Errors

ESS	Inappropriate Words	Intangible	Part of Speech	Usage	Sum
1st	17			1	18
2nd	23	2			25
3rd	15				15
4th	14	3	3		20
5th	13	2		1	15
Total	82	7	3	2	93

4.1.2 Research Question 2. Development of Accuracy, Complexity, and Fluency

A repeated measure MANOVA, using measures of learner accuracy (EFT) as dependent variables, did not show a significant effect for accuracy.

Table 3

Descriptive Statistics for Learner Accuracy (N=18)

	%EFT		95% CI	
	M (SD)	SEM	LO	HI
First	50.9 (29.0)	6.84	36.5	65.3
Second	42.8 (30.2)	7.13	27.8	57.9
Third	46.7 (21.4)	5.05	36.0	57.3
Fourth	43.1 (26.4)	6.22	29.9	56.2
Fifth	52.0 (24.1)	5.69	39.9	64.0

Note. %EFT = Percentage of error-free t-units; SEM = Standard error of the mean

Table 4

Result of Mauchly's Sphericity Test

					εb		
	Mauchly'W	X^2	F	Probability	Greenhouse-Geisser	Huynh-Feldt	LO
%EFT	.444	12.518	9	.188	.749	.927	.25

Table 5

Descriptive Statistics for Learner Complexity on SPT (N=18)

	SPT M (SD)	SEM	95% CI LO	95% CI HI
First	1.6 (.61)	.14	1.3	1.9
Second	1.7 (.53)	.12	1.4	1.9
Third	1.4 (.34)	.08	1.2	1.6
Fourth	1.9 (.48)	.11	1.7	2.2
Fifth	1.8 (.49)	.11	1.5	2.0

Note. SPT = Sentence nodes per t-unit; SEM = Standard error of the mean

Table 6

Result of Mauchly's Sphericity Test

					εb		
	Mauchly'W	X^2	F	Probability	Greenhouse-Geisser	Huynh-Feldt	LO
SPT	.706	5.377	9	.802	.880	1.00	.25

Table 7

Descriptive Statistics for Learner Complexity on TTR (N=18)

	TTR M (SD)	SEM	95% CI LO	95% CI HI
First	32.1 (5.0)	.18	29.6	34.6
Second	34.3 (5.2)	.24	31.7	36.9
Third	34.7 (4.5)	.08	32.4	37.0
Fourth	34.5 (4.0)	.94	32.5	36.4
Fifth	34.3 (5.4)	.28	31.6	37.0

Note. TTR = Type token ratio; SEM = Standard error of the mean

Table 8

Result of Mauchly's Sphericity Test

					εb		
	Mauchly'W	X^2	F	Probability	Greenhouse-Geisser	Huynh-Feldt	LO
TTR	.345	16.390	9	.06	.755	.936	.25

Table 5 and Table 7 show the results for learner complexity. However, they do not show the significant effects. The rate of sentence nodes per t-unit (SPT) and type token ration (TTR) remains consistent.

Table 9

Descriptive Statistics for Learner Fluency on WEFT (N=18)

	WEFT M (SD)	SEM	95% CI LO	95% CI HI
First	5.6 (2.4)	.57	4.46	6.91
Second	5.8 (3.1)	.74	4.26	7.42
Third	6.5 (1.3)	.31	5.86	7.20
Fourth	7.0 (3.3)	.78	5.39	8.70
Fifth	7.8 (2.6)	.62	6.53	9.16

Note. WEFT = Words error free t-units; SEM = Standard error of the mean

Table 10

Result of Mauchly's Sphericity Test

					εb		
	Mauchly'W	X^2	F	Probability	Greenhouse-Geisser	Huynh-Feldt	LO
WEFT	.362	15.658	9	.076	.663	.797	.25

Table 9 shows the result of words error free t-units (WEFT). Although the numbers increased, the effects were not significant.

5. Discussion

The hypothesis that learners will write more accurately, grammatically and lexically complex sentences, as well as more fluently on the sequence of writing tasks, was not confirmed. This result may be due to the limitations of the 50-word limit, which might make it difficult to see the developments in TTR.

The present study has some limitations that should be acknowledged. First, the

number of trials was very small. It can be difficult to find a significant difference between the results of five ESS. Second, it is not clear whether the research period was sufficiently long, as the first ESS was submitted at the end of April and the fifth ESS was submitted at the end of July. In Japan, learners need not use English outside of the classroom, known as the English as a foreign language context (EFL). It would have been useful to administer a questionnaire to students regarding how long they studied outside the classroom. Such an inquiry into this as well as the participants' approach to the assignments would have increased the study's accuracy and complexity. While some participants tried to stretch their English knowledge and made errors, others wrote with plain English. The former participants may have been very motivated, being willing to risk mistakes. Third, it is difficult to manipulate the participants' genres and control for genre effects. For example, some participants wrote a fiction story in the first writing, but they might have written poetry in the fifth writing.

The measures and evaluation also should be taken into consideration. The present study measured accuracy, complexity and fluency. These measurements are based on grammar. However, the insights of pragmatics, logicality and relevance of contexts also should be evaluated. Some students inserted an irrelevant sentence in order to meet the fifty-word rule. Other students also employed strange expressions such as *a bad face* when meaning to say "frown" and *on the night of today* for "tonight." While learners need to follow the word rule, they also need to attend to using appropriate expressions.

These findings of frequent errors are very instructive, highlighting the need for students to pay more attention to nouns, as these were the source of many errors. For example, the participants forgot to include the article or plural s with a noun. These errors were categorized as article, plural. The pronoun error was also related to numbers. When referring to plural items, the singular was used. Such errors were classified as a linguistic error (e.g., a grammatical error or an error in vocabulary use), but many of them can be understood from the context. The learners should pay

attention to content such as logicality and cohesion. These findings will be helpful in the instructive and evaluation domains.

References

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*, 191-205.

Cabinet Office. (2012). *Summary and conclusion of discussions on promoting the development of global human resources*. Retrieved from http://www.kantei.go.jp/jp/singi/global/1206011matome.pdf

Educational Testing Service (2017). *Test and score data summary for TOEFL iBT tests January 2016-December 2016 test data*. Retrieved from http://www.ets.org/s/toefl/pdf/94227_unlweb.pdf

Hassall, P. (2006). International collaboration to promote literature and linguistic via the extremely short story competition [ESSC]. *Asian Englishes*, *8*(2):pp. 90-95.

Honna, N. (2008). *English as a multicultural language in Asian Contexts: Issues and ideas*. Japan: Kurosio Publishers.

Honna, N., Kirkpatrick, A., & Takeshita, T. (2018). Across cultures: For better English communication and understanding. Japan: Sanshusha Publishers.

Lee, I. (2005).Error correction in the L2 writing classroom: What do students think? *TESL Canada Journal*, *22*, 1-16.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*, 1-40.

Ministry of Education, Culture, Sports, Science and Technology. (2013). *The course of foreign languages of English in high schools*. Retrieved from http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2010/01/29/1282000_9.pdf

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

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, *45*(1), 99-140.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, *41*, 255-283.

Van Beuningen, C., DeJong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, *62*, 1-41.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy & complexity*. Hawaii: University of Hawaii Press.

Appendices

Appendix A: Example of Errors (The error is italicized by the author.)

Agreement: They *has* a short life.

Article: It was happy week. (Definite article is dropped.)

Conjunction: *If* it is summer, I go to the sea.

Fragment: For example, astronauts, superheroes, and Hollywood stars.

Null subject: Because look at the strawberry moon. (Subject is dropped.)

Plural: There are five *womans* in the room.

Preposition: The person is kind *for* anyone.

Pronoun: But *it* already couldn't eat.

Relative clause: The close friend who plays and talks together speak from the my heart. (An error was made with the relative clause.)

Singular: So I am strong weed*s*.

Tense: Yesterday, I *drink* cafe ore.

3rd singular s: My mother always *look* displeased.

Voice: The mechanism still *don't solve*.

Word order: That *really* was too bad.

Appendix B: Example of Lexical Errors (The error is italicized by the author.)

Inappropriate words: Even if *my leg hurt*, I must dance.

Intangible: When seeing old building the left and when seeing the right it's a town of the near future.

Parts of speech: Money is necessary to *survival*.

Usage: We may be able to meet *with* entertainers.