Modern English Education, Vol. 20, No. 3, Fall 2019 (69-76)

Modern English Education

http://dx.doi.org/10.18095/meeso.2019.20.3.69

Effects of Error Correction-Prompted Revision on Learning L2 Grammatical Structures of Different Complexities

Sung-Soo Jang*

Inje University

ARTICLE INFO

Received 14 June 2019
Revised 23 July 2019
Accepted 14 August 2019

Examples in: English
Applicable Languages: English
Applicable Levels: Tertiary

KEYWORD

revision/
corrective feedback/
grammatical complexity/
수정/
교정피드백/
문법적 복잡성

ABSTRACT

Jang, Sung-Soo. (2019). Effects of error correction-prompted revision on learning L2 grammatical structures of different complexities. *Modern English Education, 20*(3), 69-76.

The study compares the effects of revision undertaken after written corrective feedback (CF) on learning two grammatical structures that differ in complexity: the articles and the hypothetical conditional. In an experiment designed for this purpose, ninety EFL college learners comprised one control and four experimental groups that were distinguished by the target structure of written CF they received and whether to carry out revision. The participants engaged in four dictogloss tasks to produce written texts. The first, second, and fourth dicotogloss writings were used as the pretest, the immediate posttest, and the delayed posttest, respectively. Learners' knowledge of the grammatical structures was measured by calculating the accuracy of using them in their writing. The results indicated that revision enhanced knowledge of the articles in the short term while it did not increase that of the hypothetical conditional. The findings suggest that the efficacy of revision on learning a particular structure is mediated by its grammatical complexity.

I. INTRODUCTION

Written corrective feedback (CF) has constituted one of the central issues in second language acquisition (SLA) literature for decades. A number of studies reported that written CF is beneficial for the development of second language (L2) knowledge (Bitchener, 2008; Bitchener & Knoch, 2010; Chandler, 2003; S. Cho, 2018; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007; Van Beuningen, De Jong, & Kuiken, 2012). It was also found that such positive effects are mediated by a variety of contextual factors (e.g., written CF type) and learner factors (e.g., anxiety, language proficiency, and language aptitude). Recently, revision made after written CF has received attention as a mediating factor.

Revision has mainly been used as a tool to measure learners' level of L2 knowledge. Some researchers, however, cast doubt on its validity as a measuring tool (Ashwell, 2000; Sachs & Polio, 2007; Truscott, 1996, 1999, 2007; Truscott & Hsu, 2008). Others claimed that revision can improve learners' knowledge, rather than simply measure it, increasing the efficacy of its preceding written CF (Ferris, 2004, 2006; Sachs & Polio, 2007). In fact, some studies provided empirical evidence to support this. (S. S. Jang, 2012, 2016; Shintani, Ellis, & Suzuki, 2014). However, the

* Author: Sung-Soo Jang (Inje University, Professor)

Sung-Soo Jang
Department of Liberal Arts Education, Inje University, 197 Inje-ro, Gimhae-si, Gyeongsangnam-do, 50834, Korea
Tel: (055) 320-3561 / Email: elansing@hanmail.net

effect of revision was shown to be affected by the type of its prior written CF (S. S. Jang, 2012, 2016). There might be a variety of other influencing factors that have yet to be empirically explored to inform how revision works for L2 learning. Among others, the study sought to investigate grammatical complexity of target structures for written CF. For this purpose, two grammatical structures of different complexity were compared in terms of the effects of revision on increasing knowledge of them.

II. THEORETICAL BACKGROUND

1. Written CF and L2 Acquisition

SLA researchers predominantly focused their attention on oral CF rather than written CF until Truscott's influential 1996 study came out. Looking seriously into written CF from the perspective of SLA, Truscott (1996) asserted that it is not facilitative but detrimental to the long-term L2 development and concluded that it "has no place in writing classes and should be abandoned" (p. 361). A number of subsequent studies, however, have produced empirical evidence that, countering Truscott's claim, written CF has beneficial effects on L2 learning (Bitchener, 2008; Chandler, 2003; Ellis et al., 2008; Sheen, 2007; Van Beuningen et al., 2012).

Another line of research has found diverse learner and contextual factors to mediate the efficacy of written CF. Sheen (2007) showed learners' language aptitude to positively influence the effectiveness of direct written CF accompanied by metalinguistic information. S. S. Jang (2013) found that writing anxiety negatively affected effects of written CF. Type of written CF has been identified as a strong contextual factor although results have been mixed as to which type is more effective. Comparing direct CF with indirect CF, for instance, some studies showed that direct CF was more facilitative (Bitchener & Knoch, 2010; Ellis et al., 2008; Van Beuningen et al., 2012) whereas other studies reported that indirect CF was superior (Ferris, 2002, 2006).

2. Revision Following Written CF in L2 Acquisition

A lot of research has used the accuracy of students' written texts revised after receiving written CF to assess change in L2 knowledge due to written CF and consistently shown that students' error has decreased in revision (Chandler, 2003, 2004; Ferris, 1999, 2003; Sheppard, 1992). Some researchers, however, did not recognize error reduction in revision as an adequate evidence for the growth of L2 knowledge on the grounds that error can be evaded during revision by using knowledge-irrelevant task strategies (Ashwell, 2000; Truscott, 1996, 1999, 2007). In Sachs and Polio's (2007) study, for example, learners improved their performance in revision by means of mechanically memorizing and simply visualizing corrections on their prior writing. Truscott and Hsu (2008) argued that progress in L2 knowledge should be evaluated through "a comparison between accuracy on the initial writing and that on a new writing task" (p. 295). The logic behind it is that genuinely acquired new knowledge is to be transferred and applied to different contexts. In fact, learners in Truscott and Hsu's (2008) study failed to produce a new piece of writing at the same level of accuracy as their revised text.

Some researchers posited that the role of post-CF revision may be to enhance L2 knowledge, rather than measure the enhanced L2 knowledge (Ferris, 2004, 2006; Sachs & Polio, 2007). They argued that revision may raise the efficacy of its preceding written CF as it provides additional opportunities for learning processes, which could lead to increasing CF-promoted gains in L2 knowledge. This view is theoretically grounded in Swain's (1985, 1995, 2005) output hypothesis, which proposes the necessity of pushing learners to produce comprehensible output for L2 acquisition. As learners seek to elevate the comprehensibility of their original text by referring to written CF during revision, it can be legitimately regarded as production of comprehensible output. In this sense, as Ferris (2010) pointed out, "revision and editing of a specific text after receiving written CF may be a helpful and perhaps necessary intermediate step between expert feedback about a target feature and long-term acquisition of that feature" (p. 189). Some studies have reported evidence for acquisition-facilitating effects of revision. Shintani et al. (2014) and S. S. Jang (2012) found that direct written CF was more beneficial to learners who went through revision and S. S. Jang (2016) showed that revision raised the efficacy of metalinguistic CF.

3. Complexity of Grammatical Structures

Research on written CF has found that it does not aid the learning of all linguistic structures (Bitchener, Young, & Cameron, 2005; Frear, 2012). Ferris (1999) distinguished structures that are used "in a patterned, rule-governed way" (e.g., regular past tense and articles) from those that have "no handbook or set of rules students can consult" (p. 6) and argued that only errors in the former are treatable by means of written CF. Supporting this claim, Frear (2012) reported that written CF worked for regular past tense but not for irregular past tense. Bitchener et al. (2005) also showed that written CF was effective for regular past tense and definite article but not for prepositions whose usage is more idiosyncratic.

It should be noted that rule-governed structures differ widely in their complexity, which is primarily regulated by the number of rules involved in usage. Complexity of a linguistic structure was shown to affect the efficacy of written CF that targeted it (S. S. Jang, 2018; Shintani et al., 2014). As such, grammatical complexity could conceivably have a mediating impact on the efficacy of post-written CF revision as well. The study aimed to explore this possibility, which little research has yet investigated. The articles and the hypothetical conditional were chosen as two target structures that notably contrast in complexity. The hypothetical conditional is regarded as more complex since it involves

representing hypotheticality as well as time reference with various forms of auxiliary and verb. Error was corrected in the form of direct CF, which is more explicit than indirect CF as direct CF provides both locus and correct form of error while indirect CF only indicates its locus. Since participants were at low-intermediate proficiency, it was deemed appropriate to offer more explicit CF. Three research questions were devised to guide the study:

1) Does written CF with revision lead to higher accuracy in learners' use of the articles than written CF without revision?
2) Does written CF with revision lead to higher accuracy in learners' use of the hypothetical conditional than written CF without revision?
3) Is there any difference in the efficacy of written CF with revision on the accuracy in learners' use of the articles and the hypothetical conditional?

III. METHOD

1. Participants

A total of 90 Korean students in three general English classes participated as part of class activities in this study. All of the participating learners were first-year university students whose majors included consulting, special needs education, biomedical chemistry, and design. Their general English proficiency was at low-intermediate level as judged by their TOEIC scores which ranged from 400 to 500. They had received formal English instruction for ten years at their primary and secondary schools.

The learners were randomly divided into four experimental groups and one control group. The experimental groups were separated from one another by two features: (a) whether the target structure for written CF the group received was the articles or the hypothetical conditional and (b) whether the group undertook revision after receiving written CF. The control group neither received written CF nor carried out revision. For brevity's sake, the experimental groups are named as AO (Article, No Revision), AR (Article Revision), HO (Hypothetical Condition, No Revision), and HR (Hypothetical Condition, Revision). The first letter of a group's name indicates the target structure for that group with A and H corresponding to the articles and the hypothetical conditional respectively. The second letter is related to whether to go through revision. O and R signify 'no revision' and 'revision' respectively. For instance, the HO group was the one that received written CF for the hypothetical conditional but did not get opportunities to revise their prior writing. The control group is named as CN. Eighteen learners were in the AO group, eighteen in the AR group, seventeen in the HO group, eighteen in the HR group, and nineteen in the CN group.

2. Writing Tasks

Four dictogloss tasks were carried out as writing activities. In each task, the participants listened to a story of some 90 words twice. They were then given a writing sheet, at the top of which a few key words from the story just heard were provided. The learners were told to recollect and reconstruct the whole story as accurately as they could for 5 minutes. Guided writing such as dictogloss was implemented since free writing seemed to be a little too difficult for the participants at the low-intermediate proficiency. It was critical to grasp and retain the content of a story for its successful reconstruction since the length of 90-odd words made mechanical and literal memorization of the whole story unfeasible. However, holding the content of a whole story in memory for 5 minutes demanded a good deal of cognitive resources. Offering some key words aimed to lessen such cognitive burden. The revision groups revised a dictogloss writing for 5 minutes immediately after receiving and studying written CF on that writing in the next task session.

The learners' first dictogloss writing constituted the pretest since it was produced before written CF and revision were implemented. Written CF and opportunities for revision were provided only to the first and second dictogloss writing. Third dictogloss writing served as the immediate posttest as it was done immediately following the final (i.e., second) written CF and revision. The last and fourth dictogloss writing which occurred two weeks later was utilized as the delayed posttest.

3. Target Structures

The articles and the hypothetical conditional constituted target structures for written CF. Each dictogloss story contained four each of indefinite article, definite article, and hypothetical conditional. Although the articles were chosen as the less complex structure, their usage is far from simple in itself. They have a variety of functions, which are embodied as a result of considering all of linguistic, semantic and pragmatic factors. In an attempt to make a striking difference in complexity between the two target structures, written CF for the articles was limitedly given on two of their functions: the indefinite article indicating that its subsequent noun is being mentioned for the first time (i.e., first mention function) and the definite article indicating that its subsequent noun has been earlier mentioned within a certain discourse (i.e., anaphoric reference function) (e.g., "He had a computer and kindly let me use the computer when mine crashed").

The hypothetical conditional is legitimately higher in complexity than the articles for the two functions. Its correct use requires not only the understanding of modal auxiliaries and tense-aspect system but the capability to encode hypotheticality and past tense reference. For example, the learner should be able to mark pastness by using the past perfect form as in "If I had given him a wake-up

call, he wouldn't have been late for the meeting."

4. Scoring Accuracy of the Target Structures

1) Articles

Accuracy in article use for each written text was obtained in percentage score using a formula for target-like use proposed by Pica (1994):

$$\frac{\text{Number of grammatical morphemes supplied accurately}}{\text{Number of obligatory contexts + Number of overused forms}} \text{X100}$$

In this study, an obligatory context indicates the case in which the use of an article was required before a noun. The number of obligatory contexts equals to that of articles included in the dictogloss passages (i.e., eight for each passage) since they were all obligatory ones. An overused form coincides with a non-obligatory article that was used in a learner's writing. A grammatical morpheme supplied accurately corresponds to any correctly used article in a learner's writing whether it was obligatory or overused.

2) Hypothetical Conditional

Accuracy for the hypothetical conditional was calculated using the scoring criteria of Shaintani et al. (2014) and Shintani and Ellis (2015) as seen in Table 1.

TABLE 1
Scoring Criteria for the Hypothetical Conditional

Clause	Criteria	Components	Point
if clause	(a)	*have* (auxiliary) + verb	1.0
	(b)	*had*	0.5
	(c)	correct form of past participle	0.5
Main clause	(d)	past modal	1.0
	(e)	*have* (auxiliary) + verb	1.0
	(f)	correct form of *have* (auxiliary)	0.5
	(g)	correct form of past participle	0.5
		Total possible	5.0

Note. The table is adapted from the one in Shintani et al. (2014, p. 115).

The study followed some of the guidelines put forth in Shaintani et al. (2014) on applying the scoring criteria:

1) The subordinate clause (i.e., if-clause) and the main clause are scored separately.
2) For the if-clause, 1 point is given if criterion (a) (i.e., the if-clause contains auxiliary have + verb) is met. Only if criterion (a) is met, 0.5 point is additionally given for satisfying each of criteria (b) and (c) (i.e., the forms of auxiliary have and past participle are correct).
3) For the main clause, 1 point is given if each of criteria (d) (i.e., the clause contains a past modal) and (e) (i.e., the clause contains auxiliary have + verb) is met. Only if criterion (e) is met, 0.5 point is additionally given for satisfying each of criteria (f) and (g) (i.e., the forms of auxiliary have and past participle are correct).

To take an example, the sentence "If I had saw the movie, I could told you about it." is scored 2.5 points. Criteria (a) and (b) are met but criterion (c) is violated since the past participle form is incorrect, which results in 1.5 points for the if-clause. Criterion (d) is met but criterion (e) is not satisfied. This leads to no point being given for criteria (f) and (g) as well as criterion (e), resulting in 1 point for the main clause.

Seventy texts (over 25% of the total) were rescored by another SLA expert to examine the scoring reliability. The interrater agreement rate was 94.5% for the article scores and 96.8% for the hypothetical conditional scores.

5. Procedure

Four experiment sessions were held over a period of four weeks during which the learners engaged in activities including dictogloss tasks, written CF inspections, and revisions. Sessions 1 to 3 occurred at five-day intervals in the first two weeks. In session 1, all the groups completed the first dictogloss task. In sessions 2 and 3, all the experimental groups received and looked over written CF on their text produced in the prior session and the revision groups (i.e., the AR and the HR groups) further made a revised text. All the experimental groups then carried out another dictogloss task with a different story. The control group went through the identical dictogloss tasks to the ones for the experimental groups without engagement with written CF or revisions. In sessions 4, which took place two weeks later, all the groups performed the last dictogloss task. Table 2 briefly summarizes the whole procedure.

TABLE 2
Whole Experiment Procedure

Session	Experimental group	Control group
Session 1	· Writing 1 (Pretest)	· Writing 1 (Pretest)
Session 2	· Examining CF on writing 1 · Revision of writing 1 [revision groups only] · Writing 2	· Writing 2
Session 3	· Examining CF on writing 2 · Revision of writing 2 [revision groups only] · Writing 3 (Immediate posttest)	· Writing 3 (Immediate posttest)
Session 4	· Writing 4 (Delayed posttest)	· Writing 4 (Delayed posttest)

IV. RESULTS

A series of factorial repeated measures ANOVAs were performed to analyze accuracy scores for the learners' written texts. The alpha level was set at $p = .05$ to decide the significance of inferential statistics. Tukey's posthoc comparions were performed to see which pairs of groups

showed significant differences. As for effect sizes, partial eta squared (η^2) was calculated for the ANOVAs and Cohen's *d* was estimated for the pairwise comparisons.

Table 3 exhibits descriptive statistics for the accuracy scores in the use of the articles for the AO, the AR, and the CN groups. A factorial repeated measures ANOVA was performed with test scores as a dependent variable and time and group as independent variables. The analysis revealed significant main effects for group, $F(2, 52) = 4.181$, $p = .021$, $\eta^2 = .139$, significant main effects for time, $F(2, 104) = 27.448$, $p < .001$, $\eta^2 = .345$, and a significant time X group interaction, $F(4, 104) = 6.338$, $p < .001$, $\eta^2 = .196$.

Tukey's posthoc analysis showed that there were no significant group differences at the pretest, which indicates that the groups had the same level of knowledge regarding the target functions of the articles at the outset of the experiment. At the immediate posttest, both experimental groups (i.e., the AO and the AR groups) performed better than the CN group. The effect sizes (i.e., Cohen's *d*) for the AO-CN group difference and the AR-CN group difference were .763 and 1.408, respectively. Comparing the experimental groups, the AR group outperformed the AO group. At the delayed posttest, both experimental groups performed better than the CN group. The Cohen's *d* values for the AO-CN group difference and the AR-CN group difference were .725 and .873, respectively. No significant difference emerged between the two experimental groups.

With regard to pairwise with-in group comparisons, both experimental groups significantly progressed from the pretest to the immediate posttest. At the delayed posttest, the AO group sustained the short-term gains while the AR group's scores significantly decreased. The CN group did not show any significant differences among the three tests.

TABLE 3
Descriptive Statistics for the Article Scores

Group	*n*	Pretest		Immediate posttest		Delayed posttest	
		M	*SD*	*M*	*SD*	*M*	*SD*
AO	18	36.48	18.83	52.59	22.89	53.02	29.34
AR	18	31.43	21.82	68.17	25.43	56.63	28.43
CN	19	30.92	17.20	36.26	19.88	33.41	24.75

Table 4 displays descriptive statistics for the accuracy scores in the use of the hypothetical conditional for the HO, the HR, and the CN groups. The analysis of a factorial repeated measures ANOVA exhibited significant main effects for group, $F(2, 51) = 4.648$, $p = .014$, $\eta^2 = .154$, significant main effects for time, $F(2, 102) = 48.535$, $p < .001$, $\eta^2 = .488$, and a significant time X group interaction, $F(4, 102) = 9.813$, $p < .001$, $\eta^2 = .278$.

Tukey's posthoc analysis revealed that no significant group differences existed at the pretest, indicating that the groups were indistinguishable in terms of knowledge of the hypothetical conditional at the beginning of the experiment. At the immediate posttest, both experimental groups (i.e., the HO and the HR groups) performed better than the CN group. Cohen's *d* values for the HO-CN group difference and the HR-CN group difference were 1.630 and .687, respectively. No significant difference was shown between the two experimental groups. At the delayed posttest, both experimental groups outperformed the CN group. Cohen's *d* values for the HO-CN group difference and the HR-CN group difference were 1.402 and .809, respectively. The difference between the experimental groups failed to reach the significant level.

As for pairwise with-in group comparisons, both experimental groups significantly improved from the pretest to the immediate posttest and sustained the short-term gains at the delayed posttest. There were no significant differences in the scores of the CN group among the three tests. Table 5 presents a summary of significant group differences at three test times.

TABLE 4
Descriptive Statistics for the Hypothetical Conditional Scores

Group	*n*	Pretest		Immediate posttest		Delayed posttest	
		M	*SD*	*M*	*SD*	*M*	*SD*
HO	17	6.62	6.13	17.00	3.53	16.09	3.84
HR	18	6.03	7.48	12.94	7.64	13.31	5.47
CN	19	6.82	5.94	7.55	8.05	7.97	7.73

TABLE 5
Summary of Significant Group Differences

	Articles	Hypothetical conditional
Pretest	AO = AR = CN	HO = HR = CN
Immediate posttest	AR > AO > CN	HO = HR > CN
Delayed posttest	AO = AR > CN	HO = HR > CN

Note. "A = B" means that there was no significant difference in scores between A and B.
"A > B" means that scores of A were significantly greater than those of B.

V. DISCUSSION

To fulfill the purpose of the study, which was to compare the impact of written CF with revision on learning two grammatical structures, it was necessary to first examine the efficacy of written CF without revision. The results showed that both the AO and the HO groups outperformed the CN group at the immediate and delayed posttest, suggesting that written CF was effective for learning the articles and hypothetical conditional alike. This could be explicated by Schmidt's (1993, 1995, 2001) noticing hypothesis, which claims that it is essential for the learner to attend to and notice a grammatical feature in order to acquire it. From this perspective, providing written CF is instrumental in L2 acquisition since it is "an attempt to draw learners' attention to problems in their writing" (Shintani et al., 2014, p. 105).

The type of written CF in the study (i.e., direct CF) may particularly have been helpful for noticing the target structures in that it provided correct form of error. A correct form was offered above each error. Such visual juxtaposition most likely made it easier for the learners

to cognitively compare and notice the gap between their flawed form and correct form of a target structure. The gap noticed this way might have led learners to pay more active attention to the target structure in the next dictogloss passage, raising their level of noticing it.

It should be noted that although written CF was effective for both target structures, the magnitude of the effects was different as measured by effect sizes. At the immediate posttest, effect sizes between the AO and the CN groups and between the HO and the CN groups were .763 and 1.630, respectively. At the delayed posttest, the same between-group effect sizes were .725 and 1.402, respectively. This shows that the effects of written CF were roughly two times greater for the hypothetical conditional than they were for articles in the short and long term alike.

The superiority of the hypothetical conditional may be ascribed to its distinctive features that could promote learners' attention and noticing in terms of saliency and semantic magnitude. For the low-intermediate learners, the hypothetical conditional was arguably more salient than articles due to its lower frequency. In fact, articles are the structure used most often in English (Sinclair, 1991) and it is most likely that they were already quite familiar to the learners. As saliency of a linguistic structure positively affects learners' noticing of it (Izumi, 2002), in all likelihood, the learners paid attention to and noticed the hypothetical conditional at a higher level than the articles.

As noted above, it was imperative for the learners to grasp and retain the content of a whole story in order to successfully carry out a dictogloss task. The hypothetical conditional in a sentence held words that carried its core meaning (e.g., verb) and mood (e.g., if and a past modal) while articles, as function words, were of relatively less semantic significance. This made hypothetical conditional more crucial for undertaking dictogloss and, as a result, it drew learners' attention more than the articles.

With regard to the effects of revision, it increased new knowledge of the articles gained from written CF treatments as the results showed that the AR group performed better than the AO group at the immediate posttest. The superiority of the AR group could be accounted for in two ways. The AR group knew beforehand that they would be required to produce a revised text immediately after written CF treatments. This may have induced the AR group to process written CF at a deeper cognitive level (e.g., pay more attention) than the AO group when inspecting it in order to enhance accuracy of their subsequent revision.

Swain (1985, 1995) contended that producing comprehensible output has three functions to help L2 acquisition processes-metalinguistic, noticing, and hypothesis testing functions. When producing output, learners cognitively "move from semantic processing to syntactic processing" (Swain, 1985, p. 249) and this "would seem to have a potentially significant role in the development of syntax and morphology" (Swain, 1995, p. 128) (i.e., metalinguistic function). Learners also get to recognize necessary forms that they need but do not know when they try to convey their intended meaning (i.e., noticing function). They further could make their own hypotheses about L2 forms in production which might be confirmed or disconfirmed through subsequent feedback (hypothesis testing function).

Revision was an act of producing comprehensible output, let alone dictogloss writing, as undertaken in this study. With literal memorization being unfeasible due to the length of a story, the leaners were made to, if in part, produce their own text to express its content (i.e., comprehensible output) in both dictogloss and revision. As such, the AR group had twice more opportunities to engage in the aforementioned output-aided L2 acquisition processes than the AO group.

The gains brought by revision, however, were not sustained over time. Unlike the immediate posttest, there was no significant difference in performance between the AO and the AR group at the delayed posttest. With-in group comparisons revealed that only the AR group's scores exhibited a significant decline from the immediate to the delayed posttest. According to skill acquisition theory, practice is required to consolidate new knowledge and convert it into long-term knowledge (DeKeyser, 1998, 2001). Practicing new knowledge refers to repeated activation of it in production for meaningful purposes. During revision, the AR group was producing a text that held the forms and content of a story as accurately as possible (i.e., production for meaningful purpose) by utilizing knowledge gained through written CF (i.e., activation of new knowledge). In this sense, revisions were legitimately regarded as acts of practice for building long-term knowledge.

Why then did the AR group fail to sustain its short-term superiority to the AO group over time? It is notable that the AR group rewrote only two times following written CF. Presumably, two-time revision was only enough to increase the gains obtained due to written CF in the short term but did not provide sufficient practice to consolidate and maintain the increased short-term gains over time. If more opportunities for revision had been provided, the short-term effects of revision could have lasted in the longer term. The issue on optimal frequency of revision for the long-term L2 acquisition has yet to be empirically addressed and certainly merits a lot of attention in SLA research.

As for the efficacy of revision for the hypothetical conditional, it had no significant impact on increasing written CF-generated knowledge either in the short or long term. As shown above, the effect size of written CF without revision was a lot larger for the hypothetical conditional. In fact, the means of the HO group at the immediate and delayed posttest were 17.00 and 16.09 out of 20.00, respectively. This indicates that the efficacy of written CF without revision was so great that there was little room for further increase in knowledge through subsequent revision, which gives an adequate explanation to why beneficial effects of revision did not emerge for the hypothetical conditional.

VI. CONCLUSION

The impact of revision on L2 learning was found to vary with complexity of its target structure. Revision was effective for increasing short-term knowledge of the less complex structure (i.e., articles) since, as an act of producing comprehensible output, it provided additional opportunities to engage in cognitive processes necessary for L2 acquisition. However, such effects were not held over time probably because frequency of revision was not high enough to allow for sufficient practice. As to the structure of higher complexity (i.e., hypothetical conditional), revision failed to improve learners' knowledge of it in the short and long term as the effects of written CF without revision were so great as to make revision futile.

The findings of the study clearly showed that revision does not work favorably for all grammatical structures and that its effects could be influenced by complexity of a target structure. As such, a teacher is advised to consider how complex a particular structure is when making decisions on whether to provide opportunities for revision if he or she is to efficiently aid students in learning that structure. On the one hand, a teacher could encourage students to revise their writing for rather simple structures. A teacher also needs to make sure students get involved in sufficient revisions to practice new knowledge if he or she wants to develop students' long-term knowledge. On the other hand, a teacher might not want to implement revision for as highly complex structures as the hypothetical conditional.

A considerable caution should be used in generalizing the findings of the study to other learners and contexts due to its limitations. The study focused only on direct CF. Given that the efficacy of indirect CF is different from that of direct CF (Bitchener & Knoch, 2010; Ellis et al., 2008; Ferris, 2002, 2006; Van Beuningen et al., 2012), effects of revision following the two types of CF might differ as well. The participating learners were all at low-intermediate proficiency level. As learners' proficiency could strongly affect the way they process and utilize written CF when writing (Ellis, 2009; Ferris & Roberts, 2001), effects of revision might vary depending on learner proficiency. Revision was undertaken two times in the study. As noted above, frequency of revision might be a factor of no little importance to determine its efficacy.

The study provided empirical evidence that, for learners, "revision is not only an act of expressing learning products but could be that of engaging in learning processes" (S. S. Jang, 2016, p. 47) although it worked only limitedly for the less complex structure. Future research should actively explore conditions for post-written CF revision to optimally engage learners in attending to, noticing, and practicing target structures so as to increase written CF-generated knowledge of them. Addressing the aforementioned limitations of the study might be a good start. And most likely, this line of research would broaden teachers' pedagogical options by raising awareness of revision as a teaching device to help learners build new knowledge beyond a measure to assess their learned knowledge.

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), 227-258.

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

Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31*(2), 193-214.

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.

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.

Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing, 13*(4), 345-348.

Cho, Sookyung. (2018). How do tutors give feedback? Case studies on their feedback on perceptions. *Modern English Education, 19*(2), 1-8.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). New York: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge: Cambridge University Press.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97-107.

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.

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), 1-10.

Ferris, D. R. (2002). *Treatment of error in second language student writing.* Ann Arbor: University of Michigan Press.

Ferris, D. R. (2003). *Response to student writing implications for second language students.* Mahwah, NJ: Lawrence Erlbaum.

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…?) *Journal of Second Language Writing, 13*(1), 49-62.

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: Contexts and issues* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*(2), 181-201.

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.

Frear, D. (2012). *The effect of written corrective feedback and revision on intermediate Chinese learners' acquisition of English.* (Unpublished doctoral dissertation). The University of Auckland, Auckland, New Zealand.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition, 24*(4), 541-577.

Jang, Sung-Soo. (2012). Effects of written corrective feedback-based revision on L2 acquisition. *Korean Journal of Applied Linguistics, 28*(4), 187-215.

Jang, Sung-Soo. (2013). Written corrective feedback and writing anxiety in L2 acquisition. *Foreign Languages Education, 20*(3), 87-112.

Jang, Sung-Soo. (2016). Revision and the effects of written corrective feedback on L2 development. *Foreign Languages Education, 23*(3), 29-53.

Jang, Sung-Soo. (2018). Does complexity of grammatical structures mediate the efficacy of written corrective feedback? *English Language Teaching, 30*(4), 69-84.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493-527.

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition, 29*(1), 67-100.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13,* 206-226.

Schmidt, R. (1995). *Attention and awareness in foreign language learning.* Honolulu: University of Hawaii Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.

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.

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal, 23*(1), 103-110.

Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System, 49,* 110-119.

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning, 64*(1), 103-131.

Sinclair, M. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255-272.

Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292-305.

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