A Study of the Effectiveness of
Written Corrective Feedback on L2 Development
by Japanese Learners of English

2020

兵庫教育大学大学院

連合学校教育学研究科

青 山 聡

A Study of the Effectiveness of
Written Corrective Feedback on L2 Development
by Japanese Learners of English

2020

Joint Graduate School (Ph.D. Program)

in the Science of School Education

Hyogo University of Teacher Education

Satoshi AOYAMA

Table of Contents

Chapter 1

Introduction

Chapter 2

Literature Review

Chapter 3

Study 1: The Effectiveness of Written CF on Text Revisions and the Writing of New Texts

Chapter 4

Study 2: The Effectiveness of Focused and Unfocused Written CF Strategies on Different

           Tasks

Chapter 5

Study 3: The Effectiveness of Written CF on the Acquisition of Explicit and Implicit Knowledge

Chapter 6

Study 4: The Effectiveness of Written CF according to Grammatical Item-Specific Proficiency Levels

Chapter 7

Study 5: Learners' Attitudes toward Written CF and Text Revisions

Chapter 8

Conclusion

Abstract

The purpose of this dissertation is to identify the most effective form of written corrective feedback (CF) for Japanese learners of English according to their English proficiency levels. In order to accomplish the purpose, the relative effectiveness of written CF was examined under some circumstances; the effectiveness of written CF on improvement in text revisions and the writing of new texts, on improvement in different types of tasks, and on development both of explicit and implicit knowledge (Studies 1 – 4). Besides, this dissertation examined learners' attitudes toward written CF, which are considered to influence the effect of written CF, in order to consider the relationship between the effectiveness of written CF and learners' affective states to written CF (Study 5).

The main findings showed that for higher proficiency learners, any type of written CF had a positive effect on L2 development. However, no predominance of any type of written CF was found. Study 1 investigated the relative effectiveness between direct written CF and metalinguistic written CF on text revisions and on new pieces of writing through three provisions of written CF, and found that written CF positively influenced the text revisions, but no clear difference was found between direct written CF and metalinguistic written CF, and that the effect of written CF on new pieces of writing was not clear. Study 2, which examined the relative effectiveness between focused direct written CF, unfocused direct written CF, and focused metalinguistic written CF, proved that no forms of written CF had any difference on improvement in accuracy examined in three different tests. Study 2 focused on the same grammatical category as Study 1, the conditionals. This is true for the results of Study 4, which examined the relative effectiveness of indirect, direct, and metalinguistic written CF in the long term, treating different grammatical category, present and past perfect tense forms. From these findings, it is possible to say that written CF is actually helpful for higher proficiency leaners in L2 development, however it is unclear as to what the most effective written CF is in this proficiency group.

On the other hand, for lower proficiency learners, metalinguistic written CF, which gives learners metalinguistic information about forms and rules, can be most effective in L2 development.

Study 1 illustrated that metalinguistic written CF had gradual positive effects on the text revisions and contributed to an increase in accuracy in the writing of new texts, while direct written CF led to improvement in accuracy only in the immediate posttest. In Study 2, focused metalinguistic written CF proved to have a long-lasting effect in two of the three types of tests. However, the predominance of metalinguistic written CF over the other types of written CF was not observed, which indicates that the effectiveness of written CF is influenced by the types of tests. In Study 4, which focused on the present and past perfect tenses, metalinguistic written CF was more effective than direct written CF only in the immediate posttest. In Study 1, metalinguistic written CF gradually improved accuracy in the text revisions and also improved it in the writing of new text accordingly. Thus, it became obvious that a single provision of metalinguistic written CF would be insufficient for certain grammatical categories.

Study 3 investigated the comparative effects of two types of written CF, direct and metalinguistic written CF strategies, on development in explicit and implicit knowledge of English present perfect tense. The findings showed that written CF had no effect on development in implicit knowledge. They also showed that for higher proficiency learners only metalinguistic written CF had immediate and long-lasting effects, while for lower proficiency learners, both metalinguistic written CF and direct written CF had immediate effects, but only the effects of metalinguistic written CF were long-lasting. These findings verified the validity of the information processing model claiming that the effects of written CF are displayed only in development in explicit knowledge.

The difference in appropriate written CF according to the proficiency level can arise from the relationship between the type of written CF and the quantity of existing explicit knowledge leaners have in long-term memory, which has a great influence on the quality of errors. Higher proficiency learners already have a significant amount of explicit knowledge, and their errors are usually caused by a lack of some small part of the knowledge or by processing failures. Irrespective of which written CF they are given, they are able to self-correct. What is needed for them is simply the information that signifies the presence of errors, which any kind of written CF tells. On the other hand, lower proficiency learners are lacking of explicit knowledge of targeted grammatical categories, and their errors are mostly caused by a lack of it. When they receive input-providing written CF, direct CF,

which provides accurate linguistic forms, they are likely to renew the information about forms and rules stored in long-term memory. It is difficult for learners to induce a correct rule needed for new pieces of writing, even if they can self-correct, using accurate forms in text revisions. When they receive output-prompting written CF, metalinguistic written CF, they are able to reform and retest hypothesis, using given metalinguistic information, and are more likely to induce correct rules used in text revisions and necessary in the writing of new texts.

The main pedagogical implications led by these findings are as follows: (i) for higher proficiency learners, any type of written CF should be provided to *mistakes*. When *errors* are present, they need metalinguistic written CF; (ii) for lower proficiency learners, metalinguistic written CF is more useful than any other CF and, therefore, should be given as many times as possible.

Acknowledgements

List of Tables

List of Figures

List of Abbreviations

ANOVA: analysis of variance

CF: corrective feedback

EIT: elicited imitation test

ESL: English as second language

ETT: English translation test

EWT: essay writing test

FL: foreign language

GJT: grammaticality judgement test

GTEC: Global Test of English Communication

L1: first language; native language

L2: second language; foreign language

M: mean

RQ: research question

SCT: sociocultural theory

SD: standard deviation

SLA: second language acquisition research

TAP: Transfer-Appropriate Processing

TOEFL: Test of English as a Foreign Language

CHAPTER 1

Introduction

1.1 Background of the Study

Feedback, which is given to students' utterances or written texts, has been considered so far as an important intervention by teachers both from the theoretical perspective among researchers and from the pedagogical or practical perspective among classroom teachers. It includes some varieties: feedback to grammatical errors, feedback to organizational errors or issues, feedback or comment to the contents of a written text, feedback to oral or written performance, even feedback to pragmatic errors and so on. Among them the feedback strategies (Ellis, 2017) that indicate to a learner that his or her output is erroneous in some way, are called *corrective feedback* (CF). It is defined as "any teacher behavior following an error that minimally attempts to inform the learner of the fact of error" (Chaudron, 1988, p. 150), and has been used as a synonym for *negative feedback* or *error treatment* in second language acquisition research (SLA). CF can be provided both orally and in a written manner, and in response to a wide range of linguistic errors.

One of the reasons why CF has been theoretically paid much attention to is its role as *negative evidence* in second language (L2) acquisition. It has been said that *positive evidence*, which shows an acceptable usage, is essential for both first language (L1) and L2 acquisition. However, it is not clear as to whether negative evidence including CF, which tells incorrectness of an utterance, is also necessary for language acquisition. Another reason why CF has been an attractive subject in SLA is the role of CF that leads to learner's *noticing* or *hypothesis formation and testing,* which are considered to be important for L2 development. Recently, the main research interest has shifted to examining the relative effectiveness of various CF strategies on L2 development in particular types of learners and situations.

Practically, CF has been paid much attention to by teachers who are struggling with various kinds of errors made by their students in their classroom, and trying to clarify whether they should correct these errors and if so, when and how. Actually, a number of teachers correct errors, using

many kinds of CF strategies every day without the firm conviction that their error corrections surely benefit learners' oral or written performance. Therefore, irrespective of whether it is offered orally or in a written manner, the findings of studies on CF have been valuable and suggestive for both SLA researchers and classroom teachers.

1.2 Focus on Written CF

CF can be divided into two primary forms: CF that is orally given, and CF that is given in a written form. In this dissertation, the main focus is placed on 'written' CF. Written CF has been a traditionally popular pedagogical practice, but relatively ignored in terms of its contribution to L2 development. This is because oral communication is more likely to draw on the learner's implicit, automatized knowledge, and therefore to be a potentially more reliable indicator of what the learner has acquired. However, this does not necessarily mean that, in terms of feedback, oral CF is any more effective than written CF. Written CF would be better able to help learners develop their explicit, conscious knowledge of the L2, thanks to its explicitness, which promotes 'noticing', to permanence of a text, which reduces the burden of the working memory capacity, and to affective comfortability, with which 'hypothesis testing' is fostered. Learners feel more comfortable undertaking hypothesis testing in written modality than in more public settings where issues of face and identity may be more threatened if hypotheses prove to be incorrect (Bitcherner & Storch, 2016). In addition, explicit knowledge can be converted to implicit knowledge as a result of practice that is appropriately contextualized according to the skill acquisition theories of Anderson and Mclaughlin (Anderson, 1993; McLaughlin, 1990).

Empirical studies have shown that providing learners with written CF has a beneficial effect on their written accuracy (Bitchener, 2008; Bitchener & Knock, 2008; Guo, 2015; Sheen, 2007; Shintani & Ellis, 2013). However, we have to be careful about in what situation the effect emerges. For example, we have to clarify whether positive effects of written CF are observed on both new texts of writing and text revisions, on either of them, or on neither of them. We are also not sure as to whether written CF leads to more or less development in implicit knowledge, or only to development in explicit knowledge when an increase in written accuracy after providing written CF

is observed. In addition, when analyzing the effects of written CF on L2 development, we must not forget the fact that individual learners, receivers of CF, differ with each other in their cognitive abilities and affective attitudes to language learning including the reactions to written CF. These cognitive and affective factors are considered to have the potential to influence the effectiveness of written CF.

Practically, teachers are very concerned about the amount of time they spend correcting the written errors that their students make, and about whether this practice is likely to benefit their learners' improvement in their original drafts and L2 development. Of course, teachers use various oral CF strategies to errors emerging in learners' utterances during classroom activities. However, the amount of time and opportunities to offer CF for each learner are limited in oral contexts. It can be assumed that in the whole-class instruction usually adopted in Japan, where students at any level of proficiency learn English in one classroom, written CF provided to their written texts would give individual learners a good opportunity for them to take individually well-tuned instruction. With these theoretical and practical aspects in mind, not oral CF but written CF is mainly focused in this dissertation.

Oral CF studies developed complex typologies of feedback strategies, which sometimes make understanding of CF difficult. Therefore, researchers have settled on the simpler typology that is based on two key dimensions – whether a strategy is input providing or output prompting, and whether a strategy is explicit or implicit (Sheen & Ellis, 2011; Lyster, Saito, & Sato, 2013). On the other hand, written CF studies have developed various taxonomies according to the aim of the research. For instance, some studies adopted the distinction between direct and indirect written CF (including metalinguistic written CF), while others adopted the distinction between focused and unfocused written CF. Oral CF and written CF have been separately studied, and therefore it is understandable that they have developed different taxonomies so far, even though they have some common features. Thus, in order to comprehensively understand what CF strategies are, a mixed-typology that was newly developed for this dissertation is introduced after each traditional typology of oral CF and written CF is explained individually.

The next chapter begins with an introduction of taxonomies of oral and written CF with reference to distinguishing features of each CF as well as some common features of CF strategies. This is followed by a discussion of the potential contribution of oral and written CF to L2

development, using a newly developed cognitive processing model with reference to the computational framework developed by Gass (1997). The main findings in the previous studies on the relative effectiveness of CF are previewed, and the problems are pointed out, focusing only on written CF. After that, the aim of this dissertation is clearly stated, and the chapter closes with an outline of the structure and focus of the following chapters, briefly introducing five discrete studies developed to accomplish the above stated purpose.

CHAPTER 2

Literature Review

2.1 Typology of CF

2.1.1 Classification of Oral CF

Early studies on oral CF were descriptive in that they focused on classifying or labelling CF. Lyster and Ranta (1997) classified oral CF into six categories, depending on their detailed observations of corrective strategies that teachers actually provided during lessons: (1) *recasts*, (2) *explicit correction*, (3) *clarification requests*, (4) *metalinguistic feedback*, (5) *elicitation*, and (6) *repetition*.

(1) Recasts

Recasts refer to the teacher's reformulation of all or part of a learner's utterance, minus the error.

> A: Traveling is much harder in those days than we might think.
> B: Oh, traveling was much harder in those days.

(2) Explicit correction

Explicit correction is the explicit and clear provision of the correct form indicating what the learner had said was incorrect. It often includes phrases such as "Oh, you mean,..." and "You should say...."

> A: Traveling is much harder in those days than we might think.
> B: No, you should say "traveling was much harder in those days."

(3) Clarification requests

Clarification requests indicate learners that their utterance has been misunderstood by the teacher or that the utterance is ill-formed in some points and that a repetition or a reformulation is needed. A clarification request includes phrases such as "Pardon me" or "What do you mean by X?"

A: I go to the hospital two days ago.
B: Pardon?

(4) Metalinguistic feedback

Metalinguistic feedback refers to comments or questions related to the error in the learner's utterance, without explicitly providing the correct form. The comments often entail the indication that there is an error somewhere. Metalinguistic information provides some grammatical metalinguistic information. Metalinguistic questions point to the nature of the error but attempt to elicit the information from the learner.

A: I go to the hospital two days ago.
B: No, it's past tense.

(5) Elicitation

Elicitation refers to some techniques that teachers depend on to directly elicit the correct form from the learner. According to Lyster and Ranta (1997), teachers elicit the completion of utterance by strategically pausing to allow students to 'fill in the blank' as it were. Or teachers ask some questions to elicit correct forms, or occasionally ask them to reformulate their utterance.

A: If it will be fine tomorrow, shall we go out for lunch?
B: If it...., if it...

(6) Repetition

Repetition refers to the teacher's repeated utterance, in isolation, of the learner's erroneous utterance with some changes of intonation so as to highlight the error.

A: If it will be fine tomorrow, shall we go out for lunch?

B: IF it WILL be fine TOMORROW?

This classification of oral CF strategies was, in a sense, complex, and therefore the simpler typology was developed based on two key dimensions - whether a strategy is implicit or explicit and whether a strategy is input providing or output prompting (Ellis, 2017). CF can be considered as a kind of input, and is usually classified as reactive (i.e., occurring after an actual error). Then, if CF can be considered as reactive negative evidence, it can be explicit or implicit (Gass, 1997). Oral CF is either explicit or implicit. Another way to distinguish one from the other is based on the provision of an accurate form for each error: input providing (i.e., provides learners with a correct linguistic form) or output prompting (i.e., pushes the learners to self-correct without a correct linguistic form). That is, oral input-providing CF, such as explicit corrections, provides learners not only with information telling that errors were made, but also with information telling correct linguistic forms for each error, whereas oral output-prompting CF, such as clarification requests, offers learners merely with information about the presence of an error. Metalinguistic CF strategies in oral or written contexts provide learners with metalinguistic information about linguistic forms and rules as well as information about the presence of errors.

These two dimensions of CF are theoretically motivated. If L2 acquisition is seen as input driven, input-providing CF strategies are to be preferred. However, if actually producing a correct form is seen as assisting acquisition, then output-prompting CF strategies are preferable. Output-prompting CF was once called as *negotiation of meaning* (Lyster, 1998; Mackey, Gass & McDonough, 2000), but now it is sometimes called as *prompts*, which "include a variety of signals, other than alternative reformulations, that push learners to self-repair" (Ranta & Lyster, 2007, p. 152). The choice of implicit or explicit CF strategies depends on the importance of conscious noticing of the correction. Implicit CF caters to implicit acquisition, whereas explicit CF is more likely to lead to conscious noticing and explicit learning. In Lyster and Ranta's taxonomy, recasts and explicit correction are considered to be input-providing CF, but recasts are more implicit and explicit correction is less implicit (more explicit). On the other hand, clarification requests, metalinguistic feedback, elicitation and repetition are regarded to be output-prompting CF, and among them, clarification requests and repetitions are more implicit than metalinguistic feedback

and elicitation. It should be noted that explicit and implicit strategies are not two discrete components. Rather, they lie in a single continuous component.

2.1.2 Classification of Written CF

Since the 1990s, studies on written CF have been conducted, following the flourishing studies on oral CF. Written CF is different from *feedback on writing* in that the latter includes any comment on the contents, and it is given such a definition that "a written response to a linguistic error that has been made in the writing of a text by an L2 learner" (Bitchener & Storch, 2016, p. 1).

Written CF is generally 'explicit' in the sense that its corrective force is overt to a learner mainly due to permanence of a text where there are, for example, some underlines or acceptable forms next to errors. Thus, written CF strategies are often divided into two types on the basis of the provision of information about a correct linguistic form for each error, in addition to information about the presence of an error: direct written CF and indirect written CF. Direct written CF is an input-providing strategy that directly offers an accurate linguistic form near an error on a handout. On the other hand, indirect written CF is an output-prompting strategy that only indicates the presence of an error without any accurate linguistic form, and encourages learners to self-correct by means of, for instance, highlighting errors by underlining them or leaving the total number of errors on a handout (Bitchener & Storch, 2016).

In addition to these two types of written CF, there have been recent studies into metalinguistic written CF. It is defined as "that which provides that learner with an explanation of what has caused the error (and often this is in the form of grammar rules) and examples of correct usage. This is usually done by giving each error a number and at the bottom of the page of text or at the end of the full text providing the metalinguistic explanation and example(s) beside the relevant number assigned to the error category in the learner's text" (p. 17). Metalinguistic written CF strategies consist of metalinguistic information about grammatical rules and sometimes linguistic forms that are used to explain the rules as well as information telling the presence of errors. However, they do not provide a correct linguistic form itself for each error. Linguistic forms in metalinguistic written CF appeared in an example or explanatory sentence are used by learners to better understand the relevant grammatical rules.

2.1.3 A Mixed-Taxonomy of Oral and Written CF

With regard to the classification for both oral and written CF, the way in which CF is offered is sometimes used to characterize it as well as the properties of CF, such as the explicitness or the presence of a correct form for each error. That is, written CF can be also classified, following these dimensions: focused or unfocused, immediate or delayed, and single-provision or multi-provisions.

The question of how many linguistic categories CF should focus on at one time has attracted pedagogical interest among teachers. Focused written CF is given to errors on some specific linguistic categories. Feedback on only one targeted category of error is called 'highly focused' CF, while feedback on a limited number of targeted error categories is called 'less focused' CF (Ellis, Sheen, Murakami, & Takashima, 2008, p. 356). On the other hand, unfocused CF, or comprehensive CF, refers to feedback given on a wide range of error categories.

The difference in the timing of giving feedback is also used to distinguish CF strategies. Immediate CF is feedback provided immediately after the emergence of an error, while delayed CF is feedback provided after an activity was completed. Written CF is more or less invariably delayed, as it is provided after learners have completed a piece of writing (Li, Zhu, & Ellis, 2016).

Moreover, based on the frequency of CF treatment, CF can be classified into short-term treatment or long-term treatment. Short-term treatment of CF refers to a more focused approach providing learners with CF on a single occasion, even including one-off provision of CF. Long-term treatment of CF refers to an approach where an opportunity of giving CF is set in multiple occasions with some intervals of time.

Table 2.1

A Classification of CF

| | Manner | Category | Timing | Frequency | Explicitness | |
					Implicit	Explicit
Input providing						
	Oral	Focused/Unfocused	Immediate/Delayed	Single/Multi	Recasts	Explicit correction
	Written	Focused/Unfocused	Delayed	Single/Multi	—	Direct CF
Output prompting						
	Oral	Focused/Unfocused	Immediate/Delayed	Single/Multi	Clarification Requests, Repetitions	Metalinguistic feedback, Elicitation
	Written	Focused/Unfocused	Delayed	Single/Multi	—	Indirect CF, Metalinguistic CF

Therefore, to be exact, each CF can be characterized on the basis of six indexes: (1) manner (oral or written), (2) explicitness (explicit or implicit), (3) the provision of correct linguistic forms (input-providing or output-prompting), (4) the number of targeted linguistic categories (focused or unfocused), (5) the timing of feedback (immediate or delayed), and (6) the frequency of CF provision (a single provision or multiple provisions) (Table 2.1).

2.2 L2 Development through CF

2.2.1 Defining the Term 'L2 Development'

It is generally an accepted idea that the goal of L2 development is to acquire communicative competence. Communicative competence is a term coined by Hymes (1972), and consists of four components: linguistic, sociolinguistic, discourse, and strategic competence. A language learner needs to use the language not only correctly (mainly based on linguistic competence), but also appropriately (based on other three competence). What 'L2 development' means should originally include the balanced development in each of the four competencies. One of the components, the linguistic component, includes the knowledge of the sounds and their pronunciation (i.e., phonetics), the rules that govern sound interactions and patterns (i.e., phonology), the formation of words by means of inflection and derivation (i.e., morphology), the rules that govern the combination of words and phrases to structure sentences (i.e., syntax), and the way that meaning is conveyed through language (i.e., semantics). Written CF is typically given to errors in grammar and vocabulary, so it can be considered to particularly influence development in the 'linguistic' competence among four components of communicative competence, which is usually examined by an increase in accuracy in a written text.

In terms of linguistic knowledge that characterizes the linguistic competency related to accuracy in language use, two types have been identified: implicit knowledge and explicit knowledge. Implicit knowledge is the type of knowledge used automatically and with no conscious attention. Explicit knowledge is, on the other hand, used with a controlled and conscious attention to target-like accuracy. Taken together, in this dissertation, L2 development refers to development in linguistic knowledge measured on the basis of an increase in written accuracy that is led by the

acquisition of both explicit knowledge and implicit knowledge or of either of them.

The construct of L2 development is sometimes used interchangeably with L2 learning and L2 acquisition. L2 learning and L2 development are most often used interchangeably to refer to the process or processes of learning from the learner's perspective, even though the term L2 development is, arguably, more about specific stages in the learning process. L2 acquisition can be understood in terms of the acquired end-product with which learners can use the target language automatically and without conscious attention. In this dissertation, the term L2 development is used because it would be a more precise term that includes reference to any or all of the stages in L2 development, from the initial CF input stage to the implicit, automatized output stage.

Here, the key question is whether or not CF triggers development in linguistic competence, and if CF actually triggers, whether or not the development in linguistic competence is caused by development in both of two types of knowledge, or in either of them.

2.2.2 Information Processing for L2 Development through CF

In framing the discussion of the information processing in a single CF episode, the computational framework for a model of second language acquisition developed by Gass (1997) is mainly drawn upon because "the model ... constitutes the fullest and clearest statement of the roles played by input and interaction in L2 acquisition currently available" (Ellis, 2008, p. 268). The model progresses according to five main stages in the cognitive processing of input to output: (1) a*pperceived input (apperception)*, (2) *comprehended input (comprehension)*, (3) *intake*, (4) *integration* and (5) *output*.

At the first stage, *apperception*, the learner needs to apperceive or notice the gap in his or her L2 knowledge. For this to occur, the learner needs to consciously attend to the input that has been provided. As Schmidt (1990, 1994, 2001) explains, there are three levels of attention: (1) *alertness*, which explains the learner's motivation and readiness to learn, (2) *orientation*, which refers to the learner's attention to linguistic forms or accuracy, not only to meanings, and (3) *detection*, which refers to the cognitive registration of input being present for the processing of information.

The second stage of the framework, *comprehension*, explains the importance for input to be comprehended before it can become intake (Stage 3). As widely known, comprehended input is not the same as comprehensible input (Long, 1981, 1996). Comprehended input explains whether or

not the learner has actually comprehended the input.

The third stage, *intake*, requires the learner to match the input with each existing knowledge. The matching processing contains different levels of analysis in the working memory capacity comparing between the learner's existing knowledge in the long-term memory, and the input that has been received. During the process of comparison, the learner makes hypotheses about what is acceptable and what is not acceptable in the L2.

As each hypothesis is tested by means of a modification to the learner's original output, any one of four outcomes is possible in the process of the fourth stage, *integration*. First, the learner's existing L2 hypothesis, drawn from knowledge stored in long-term memory, will be either confirmed or rejected. Second, the learner's current hypothesis will be strengthened through a confirmation of the accuracy of a new use of the linguistic item. The third possible outcome is storage. The information in this input is not immediately incorporated into the learner's L2 knowledge but is stored until the learner has received more evidence later. The fourth possible outcome is one in which the hypothesis may exit from the processing system because the learner realizes it is incorrect. Before output, learners have acquired some implicit and explicit knowledge in long-term memory, which are either correct or incorrect, and wait for being used in comprehension and production.

The last stage is *output*, which is the overt manifestation of whether or not the learner has begun the process of developing linguistic competence. According to Gass (1997), output may provide learners with four important functions for language learning: "testing hypotheses about the structures and meanings of the target language; receiving crucial feedback for the verification of these hypotheses; developing automaticity in interlanguage production; and forcing a shift from meaning-based processing of the second language to a syntactic mode" (pp.139-140). Considering the contribution of CF strategies, which are classified as 'reactive' negative evidence, to L2 development, this stage *output* can actually be a starting point. In Japan, grammatical rules would be explicitly taught during a lesson, focusing on one single grammatical category at one time. Through the instruction, where they experience the stages from input to integration, learners would store some degree of information about the target language in long-term memory, and acquire some degree of explicit and implicit knowledge.

If the output does reveal an accurate use of the L2, the learning process goes into the

consolidation phase where learners can develop automatic processing through output. If the output leads to CF from the interlocutor then, learners can attend to CF and notice the presence of errors in their output and mismatch or gap between their production and given correct forms. As a result, this noticing-the-gap leads to reassessment, including hypothesis reformation and retesting, which may be on the spot reassessment in the case of oral production, or longer-term complex thinking in the case of written production. The latter can be also accomplished by gathering additional information from a variety of sources. Written CF is usually a delayed strategy, while oral CF is immediate, and therefore learners can search for the information required for reassessment not only in CF but also in, for instance, a grammar book or dictionary after receiving CF. That is, learners are able to have many kinds of information resources for reassessment if they want. In the process of reassessment, in other words, the process of hypothesis reformation and retesting, learners can depend on three kinds of information resources at hand according to CF they receive: the CF-driven information about an accurate form for each error, the CF-driven metalinguistic information about forms and rules, and the existing information in their long-term memory. Renewed information and hypothesis, which result from the reassessment, also differs according to the type of CF, and stays in long-term memory waiting for a chance to be produced, or a chance to be reassessed again. If CF pushes learners to produce the renewed information, they get four benefits for language learning again. In addition, output gives a chance to notice the hole, which would result in a search for help to the interlocutor or grammar books, for instance.

To sum up, considering this cognitive model of L2 development, when learners produce something in an oral or a written mode, they obtain an opportunity to receive CF, which may lead them to notice the gap between existing linguistic knowledge in their long-term memory and the information that has been received through CF. Noticing, then, leads to immediate or delayed reassessment of hypothesis (hypothesis reformation and retesting), which leads to storage of renewed knowledge. When it is produced orally or in a written manner, learners have the benefits of testing hypotheses, receiving feedback, developing automatized production (of course, this needs a significant amount of practice), forcing a shift of meaning-based to form-based processing, and noticing the hole, all of which are considered to be important for L2 development.

Through the stages in the information processing framework, it seems that learners can obtain more explicit knowledge than implicit knowledge by means of written CF. Therefore, we can

assume that it is difficult for written CF to directly contribute to development in implicit knowledge. In order to explain the possibility of conversion from explicit knowledge into implicit knowledge, the interaction theorists argue that explicit knowledge can be converted to implicit knowledge if certain conditions are satisfied. According to Dekeyser (1998), who supports the strong interface position, explicit knowledge can be converted into implicit only through *practice* in actual communication. Practice is seen as an important term in this context. Traditionally, practice has been viewed as an activity that involves the process of repeatedly and deliberately attempting to produce a specific feature of the target language, but, according to Dekeyser, it is more important to focus on behavior rather than structure. Therefore, mechanical practicing of a linguistic feature in decontextualized activities (e.g., mechanical drills) is seen as unlikely to affect the learner's long-term memory and to lead to a change of behavior (i.e., from controlled processing to automatic processing). On the other hand, the weak interface position (N. Ellis, 2005), while also stating that explicit knowledge can be converted to implicit knowledge, explains explicit knowledge of developmental feature would only be expected to be converted if the learner was at the developmental stage required for performing them without conscious attention. Irrespective of whether the interface position is strong or weak, in order to acquire implicit knowledge, further practice of retrieving the stored knowledge and of accurately using forms or structures in contextualized situations are at least required. In other words, explicit knowledge stored in the integration stage, could become implicit only by pulling it out many times through a significant amount of practice after approval for accurate use in the stage, output.

2.2.3 A Mechanism of L2 Development through CF

In order to understand the influence of CF on L2 development, focusing on the information CF provides is useful. Both oral and written CF are divided into three types according to the information CF provides when their contribution to L2 development is considered. Input-providing strategies (Type 1) include input-providing CF providing an accurate linguistic form for each error, such as recasts and explicit correction included in oral CF strategies, and direct written CF. Output-prompting strategies with no additional information (Type 2) contain output-prompting CF providing no additional information except for the information telling the presence of some errors, such as clarification request, repetition and elicitation in oral CF strategies, and indirect written CF.

Table 2.2

Summary of Information Given by CF and Information Potentially Stored in Long-Term Memory after Reassessment

	Information given	*Information potentially stored*
Type 1: Input-providing		
Oral (REC, ExC)	Accurate linguistic form	Accurate linguistic form
Written (Direct CF)		+
		Reformed information about form and rule
Type 2: Output-prompting		
Oral (ClR, REP, ELI)	(no information)	Reformed information about form and rule
Written (Indirect CF)		
Type 3: Output-prompting		
Oral (MF)	Metalinguistic information	Metalinguistic information
Written (Metalinguistic CF)		+
		Reformed information about form and rule

Note. REC = Recasts, ExC = Explicit Correction, ClR = Clarification Request, REP = Repetitions, ELI = Elicitation, MF = Metalinguistic Feedback

Output-prompting strategies with metalinguistic information (Type 3) include output-prompting CF providing metalinguistic information related to errors, such as oral metalinguistic feedback and metalinguistic written CF (Table 2.2). It is, of course, uncertain whether learners can actually renew and store linguistic forms from the grammatical rule or the grammatical rule from linguistic forms, or whether the renewed information about the linguistic forms or the grammatical rules are really accurate.

2.2.3.1 Type 1: Input-Providing CF Strategies

Noticing the 'gap' is fostered when input-providing strategies, such as recasts and explicit correction, which offer information about correct forms as well as information about the presence of errors, are provided. According to Long's updated Interaction Hypothesis (1996) or Schmidt and Frota's Noticing Hypothesis (1986), learners can notice the gap or mismatch between correct forms or structures that oral CF shows and their existing knowledge when they are provided with recasts. However, recasts, which are classified as 'implicit', might be ignored because their corrective forces are covert. The other oral input-providing strategy, explicit correction, is more likely to be noticed because its corrective force is clear to learners. There is a danger, however. Explicit correction might ruin the learner's motivation, which may prevent him or her from initiating language learning process. Teachers are sensitive to how their utterances, including oral CF, affect learners' affective states, and therefore they are likely to prefer the implicit oral input-providing strategy, recasts, to the explicit one.

In the case of direct written CF, every strategy is explicit due to its clear corrective force, permanence of a text, and sufficient time allocated for cognitive processing. Therefore, it can be more noticeable than oral CF. The fleeting nature of oral CF might give learners less opportunity to notice. Moreover, thanks to them, less proficient learners, who possess only limited working memory capacity, can easily notice the gap in a written manner.

By means of attended and noticed input-providing CF, linguistic hypothesis related to the error is reformed and retested with the related knowledge existing in long-term memory. As a result, newly given information about the correct form for each error and renewed linguistic information about forms or rules might be stored in long-term memory in the stage of integration. The renewed information, of course, stays as hypothesis, and therefore whether it is actually correct or not is uncertain.

In a written manner, learners are provided with enough time to properly consider and to search for additional information resources outside CF in the process of reassessment. After the writing task is returned to learners with written CF, they can refer to, for example, a grammar book if they want. In this sense, written CF has more opportunity to bring accurate renewed information about forms and rules in long-term memory.

2.2.3.2 Type 2: Output-Prompting CF Strategies with No Additional Information

Output-prompting CF strategies, such as clarification requests, repetitions, elicitation and indirect written CF, provide the information about the presence of errors. Because the corrective force of oral output-prompting CF in this type is usually unclear to learners, there is a high risk that learners cannot attend to and notice the strategies. Indirect written CF is more overt than oral output-prompting CF, but less overt than direct written CF or metalinguistic written CF, and therefore there is also the possibility that learners do not notice it. Even if these strategies are noticed, it would be so difficult for learners to reform and retest hypothesis, and finally, to store correct, renewed knowledge because what learners can rely on to reform and retest it is only existing linguistic knowledge already stored in their long-term memory. The result can be, of course, that they cannot correctly renew the linguistic forms or rules. Even if they can, learners are not sure whether these forms or rules are really correct or not.

For this reason, it is possible to say that it is only when their errors are not 'errors' but 'mistakes' that learners can correct with output-prompting CF in Type 2. According to Corder (1967), 'errors' are made as a result of a lack of explicit knowledge, while 'mistakes' reflect processing failures in performance that arise, for example, as a result of the limitation in the working memory capacity. In the case of errors, learners cannot reform and retest a new hypothesis only with information about the presence of errors. They have to ask the interlocutor during conversation or refer to a grammar book in order to receive more explicit input-providing CF, for example.

The output-prompting strategies in Type 2 would encourage learners to output more strongly than input-providing CF, which can lead to notice the 'hole' (learners want to say something, but they don't know how to say in the target language). Swain (1985) proposed the Output Hypothesis, arguing that not only comprehensible input but also comprehensible output is also important for language acquisition during interactional negotiation. When learners are required to produce 'pushed output' and make output comprehensible, they usually engage in semantic and syntactic processing. Nobuyoshi and Ellis (1993) reported that when learners made errors and received CF in the form of output-prompting CF such as clarification requests, they tried to modify their output by self-correcting their errors and, subsequently, showed improved accuracy in later tasks. *Uptake* is optional output and defined as "a student's utterance that immediately follows the teacher's feedback and that constitutes a reaction in some way to the teacher's intention to draw attention to

17

some aspect of the student's initial utterance" (Lyster & Ranta, 1997, p. 49). This uptake, or pushed output by oral CF, can be thought of as equivalent to the revision of the learner's initial piece of writing or to a new piece of writing by written CF. Immediate hypothesis-testing in uptake by oral output-prompting CF or delayed hypothesis testing in revisions or new pieces of writing by written output-prompting CF optimizes the learning potential in that learners can access to not only meaning, but also syntactic processing, obtain a chance to receive another new CF while interacting with the interlocutor or the teacher to reform hypothesis, or by referring to other information resources, develop automatic processing, and notice the hole.

2.2.3.3 Type 3: Output-Prompting CF Strategies with Metalinguistic Information

The other output-prompting CF strategies include metalinguistic oral and written CF, both of which provide metalinguistic information about forms and rules. Learners cannot directly obtain information about a correct form for each error (that's why this strategy is not called 'input-providing') from these strategies, but instead, can obtain metalinguistic information such as "you should use the past tense." When receiving this type of CF, learners can reform and retest hypothesis, using newly given metalinguistic information in CF and existing explicit knowledge. Stored knowledge in long-term memory after hypothesis reforming and retesting may be metalinguistic information and renewed linguistic information about forms and rules, which would be tested in the following opportunities for output. In the case of metalinguistic written CF, learners are able to depend on outer resources of information for hypothesis reformation and retesting just like other written CF strategies.

These output-prompting strategies in Type 3 can also lead to output, which provides learners with four functions: testing hypothesis; receiving crucial feedback for the verification of these hypotheses; developing automaticity in interlanguage production; and forcing a shift from meaning-based processing of the second language to a syntactic mode. In addition, they give a chance for learners to notice the hole.

Theoretically considering the role of CF in L2 development, it turned out that CF probably contributes to development in explicit knowledge rather than implicit knowledge, and that the difference in explicitness of CF influences noticeability. Moreover, it is also probable that learners

```
Consolidation
Automatization
        ↑
        ┊
┌─────────────┐
│    No C F    │
└─────────────┘
        ┊
┌─────────────────────┐
│  Accurate output    │
└─────────────────────┘

┌──────────────────────────────────────────────────┐
│  ┌──────────┐  ┌──────────┐  ┌──────────┐          │
│  │ Type 1   │  │ Type 2   │  │ Type 3   │          │
│  │Accurate  │  │No        │  │Metaling- │          │
│  │form to   │  │additional│  │uistic    │          │
│  │each error│  │information│  │information│          │
│  └──────────┘  └──────────┘  └──────────┘          │
│                    C F                             │
└──────────────────────────────────────────────────┘

┌─────────────────────┐
│  Inaccurate output  │
└─────────────────────┘
```

Type 1	Type 2	Type 3
Accurate form to each error	No additional information	Metalinguistic information

C F

No C F

Accurate output **Inaccurate output**

OUTPUT

INTEGRATION

INTEGRATION

Accurate form + Renewed information about form and rule	Renewed information about form and rule	Metalinguistic information + Renewed information about form and rule

INTAKE / REASSESSMENT

Accurate form + Existing explicit knowledge + Some resources	Existing explicit knowledge + Some resources	Metalinguistic Information + Existing explicit knowledge + Some resources

HYPOTHESIS REFORMATION → RETESTING

INTAKE

HYPOTHESIS FORMATION → TESTING

COMPREHENDED INPUT

APPERCEIVED INPUT

NOTICING THE GAP

NOTICING THE GAP

INPUT

A model by Gass (1997)

Figure 2.1. A model of cognitive processing for L2 development through CF.

can store different kinds of information in their long-term memory, depending on the type of CF

(Figure 2.1). Thus, it is important to investigate the effects of CF on L2 development, identifying what kind of information each CF offers, what kind of information learners can renew, and what kind of information or knowledge they can store in the end.

In order for explicit knowledge to be converted into implicit knowledge, repeated retrievals of explicit knowledge from the long-term memory during meaningful practice are needed. Through them, less controlled processing changes into more rapid, automatised processing. Therefore, output-prompting CF, specifically oral output-prompting rather than written output-prompting, is more desirable to elicit numerous opportunities to retrieve explicit knowledge. Uptakes prompted by oral output-prompting CF are urged immediately after CF because the strategy is provided in the interaction with an interlocutor, and therefore it is difficult for learners to keep on conversation, intentionally ignoring it. In contrast, a feeling of pressure or motivation to output or self-correct after written output-prompting CF would be lower because learners can ignore it, which deprives explicit knowledge of an opportunity to change into implicit knowledge.

2.3 The Relative Effectiveness of CF

2.3.1 Oral CF

Researchers have investigated the relative effectiveness of oral CF on the basis of comparison between implicit and explicit feedback, and between input-providing feedback (e.g., recasts) and output-prompting feedback (e.g., prompts in the form of elicitation, clarification requests, and repetition) separately.

Ellis, Loewen, and Erlam (2006) studied the relative efficacy of implicit and explicit on learners' development in regular past tense by comparing an implicit type of CF, recasts, and an explicit type of CF, metalinguistic feedback. They found no significant effect for both CF types on the immediate posttests but found that the explicit CF group outperformed both the implicit CF group and the control group on the delayed posttest. Sheen (2007) also found that whereas explicit CF (in the form of metalinguistic feedback plus provision of the correct form) resulted in significant gains in learning in both immediate and delayed posttests, the implicit CF did not. Both of the studies above used intact intermediate level classes of adult learners, and CF was provided in the context

of communicative activity. Thus, in a communicative L2 classroom context, explicit CF seems to be more effective than implicit recasts. However, it should be noted that a number of laboratory-based studies (e.g., Han, 2002) have shown that recasts can be also effective and facilitate acquisition.

Lyster (2004) and Ammar and Spada (2006) investigated the relative effectiveness of input-providing oral CF and output-prompting oral CF. Lyster (2004) compared the effects of recasts (as an input-providing CF) and a mixture of output-prompting CF strategies on the acquisition of gender marking on articles and nouns in French. The latter group was the only group to outperform the control group on every measure. In another study, Ammar and Spada (2006) compared the effects of recasts and prompts on learning of possessive pronouns. Prompts were especially effective for learners who had pretest scores below 50 percent, whereas learners with the score above 50 percent benefited similarly from both recasts and prompts. Taken together, these studies suggest that CF that prompts learners to self-correct, that is, output-prompting CF, is more effective than CF such as recasts, at least, for learners who have already begun to acquire the target feature.

2.3.2 Written CF

The comparative studies on the relative effectiveness of written CF have generally illustrated that direct written CF is more effective than indirect (Bitchener & Ferris, 2012). Bitchener and Knoch (2010) divided the participants into four groups: a group which receives metacognitive explanation, a group which receives indirect written CF, a group which receives metalinguistic explanation and explicit instruction, a group which receives no feedback (a control group), and compared them. The results showed that three experimental groups outperformed the control group in the immediate posttest, and in the delayed posttest, two experimental groups outperformed the experimental group which received indirect written CF and the control group, which showed that only direct written CF had a long-lasting effect. Van Beuningen, De Jong, and Kuiken (2008, 2012) also found that even though there were positive short-term effects for both direct and indirect feedback, direct error correction had a more significant long-term effect than indirect written CF.

The relevant studies have examined the effectiveness of written CF on L2 development on the basis of the classification of direct written CF and indirect written CF, and therefore have not fully included metalinguistic written CF. Guo (2015) found that the Chinese EFL learners who

received more explicit types of written CF (direct error correction; metalinguistic explanation; direct error correction plus metalinguistic explanation) outperformed those who received the less explicit types of feedback (underlining and error code) and that there was no difference between the three most explicit types. Similarly, Shintani and Ellis (2013) found no significant difference between direct error correction and written metalinguistic explanation in the effectiveness in the use of the indefinite article (but at the immediate posttest, the metalinguistic explanation group outperformed the direct error correction group). On the other hand, the study by Shintani, Ellis, and Suzuki (2014) showed that direct error correction was found to be more effective than metalinguistic explanation.

With these findings in mind, it is acceptable that explicit oral CF is more effective than implicit one because learners are more likely to notice its corrective force if the strategy is clear. Whether or not learners can attend to CF is crucial for initiating the cognitive processing for L2 development. However, it is surprising and interesting that output-prompting CF such as prompts is more effective than input-providing CF such as recasts in oral contexts, while input-providing CF such as direct written CF is more effective than output-prompting CF such as indirect written CF in written contexts. This can be explained from the frequency of output that learners produce as self-correction, which leads to a syntactic processing, as explained above. Thus, in order to examine the effects of written output-prompting CF, it would be important to make sure that after the provision of written CF, learners actually self-correct their errors in their revision of the initial piece of writing.

2.4 Problems and Limitations in the Previous Studies on Written CF

In this section, focusing only on written CF treated in this dissertation, problems or limitations are stated: those related to 'effectiveness'; those related to CF types for comparison; those related to individual learner-internal differences. These are followed by an explanation of other problems related to a research design.

2.4.1 Problem of How 'Effectiveness' Is Measured
2.4.1.1 Text Revisions or the Writing of New Texts
When stating the relative effectiveness of three types of written CF, i.e., indirect, direct, and

metalinguistic written CF strategies, on L2 development, the difference between text revisions and the writing of new texts appears to hold great importance. In other words, we should distinguish the effects of written CF on improvement in text revisions, i.e., self-corrected versions of an initial writing, from the effects on improvement in the writing of new texts. This is mainly because there is a difference in information that each written CF provides as explained before. The information that direct written CF offers is about a correct linguistic form for each error, and this becomes helpful in the case of text revisions because learners can directly use the form, even if they do not understand the rule when revising their original writing. However, in the case of the writing of new texts, direct written CF strategies may not be helpful because learners have to understand why and how the form and structures are used. In the latter case, metalinguistic written CF would be the most effective, which provides metalinguistic information about forms and rules on which learners can rely in a new piece of writing. It is for this reason that the relative effectiveness of written CF must be examined both in text revisions and in new pieces of writing.

Traditionally, the effectiveness of written CF has been hotly debated on the basis of the difference between text revisions and the writing of new texts, rather than of the difference between explicit and implicit knowledge. Although most teachers assume, to some extent, that written CF contributes to the learning process in some way (Ferris, 2003), Truscott's (1996) call for the abandonment of the practice, challenging this assumption. He argued that there was no compelling research evidence of the benefits of written CF for L2 development. He also claimed that that written CF is effective not in the writing of new text but in the learners' text revisions, and that written CF even has harmful effects because learners who are corrected tend to shorten and simplify their writing so that they avoid making too many errors. More specifically, Truscott advanced three major arguments against the effects of written CF. First, he contended that there is no empirical evidence to support the claim that written CF assists L2 learners in improving their accuracy. Second, he further claimed that written CF cannot contribute to development in L2 competence or influence the natural order and sequence of second language acquisition. Third, he argued that the provision of written CF creates many practical problems ranging from the inconsistent way in which feedback is provided, students' negative attitudes toward written CF, to anxiety and a lack of motivation that written CF generates (Truscott, 1996, 1999, 2007). On the other hand, Ferris, who stands in an affirmative side, maintained that written CF can lead to improvement in learners' grammatical

accuracy when written CF is unambiguously and consistently given (Ferris, 1999, 2003). Moreover, in reaction to Truscott's claim, a number of studies have been conducted, examining not only the effects of written CF on text revisions, focusing on the role of written CF as an editing tool, but also the effects of written CF on the writing of new texts, focusing on the role of written CF as a learning tool. They have mainly illustrated the beneficial effects of written CF on new writing texts so far (Bitchener, 2008; Bitchener & Knoch, 2010; Shintani et al., 2014), however, they are sometimes criticized for overgeneralizing the effects which proved to be clear only in a limited number of linguistic categories (Xu, 2009).

Although most preceding studies have treated the effects of written CF on a text revision and those on a new piece of writing separately, a few studies investigated them in a single study and tried to reveal whether the improvement in accuracy in text revisions leads to that in the writing of new texts. Truscott and Hsu (2008) failed to illustrate it, while Van Beuningen, Jong, and Kuiken (2012) succeeded. Thus, we have been lacking the preceding studies in order to judge whether written CF has a positive effect not only on text revisions, but on the writing of new texts (Sheen, 2011). In addition, as Bitchener and Storch (2016) cautioned, we should not conclude that the studies implying that written CF is effective show that learners have reached the level of native-like competence, that is, fully gained implicit knowledge. A period of consolidation is required for learners to convert explicit knowledge (demonstrated in immediate posttests and delayed posttests) to unconsciously retrieved and used implicit knowledge (demonstrated through consistent accuracy on multiple occasions and in multiple contexts over time). Thus, the hybrid research where the effectiveness of written CF on a text revision and on a new piece of writing is examined at the same time in a single study should be conducted (Van Beuningen et al., 2012).

2.4.1.2 A Single New Writing Task or Multiple New Writing Tasks

The effects of written CF on a new piece of writing should be examined, using different types of writing task. In classroom settings in Japan, learners perform many kinds of writing tasks such as a Japanese-English translation task and an essay writing task. These tasks must differ in cognitive load on the working memory capacity, which is one of the main constructs of L2 proficiency. A writing task such as a translation task would be less cognitively demanding than a task such as an essay writing task. In addition, from a pedagogical point of view, teachers are more interested in

whether the learner's performance in various tasks, which owes to both explicit and implicit knowledge, improved with written CF, than in which type of knowledge, explicit or implicit, was actually used. For these reasons, the effectiveness of written CF even in the writing of new texts cannot be estimated only by a single writing task. However, we are lacking the studies on the effectiveness of written CF on improvement in different kinds of tasks that are conducted in a single research.

2.4.1.3 Explicit Knowledge or Implicit Knowledge

When stating the relative effectiveness of CF strategies, we also need to compare the direct effects of written CF strategies on the acquisition of explicit knowledge, or on that of implicit knowledge. The main purpose of this separation is to examine whether negative evidence including CF actually is essential for learners' L2 acquisition. In this sense, this question is theoretically motivated.

Polio (2012) suggested that written CF leads to improvement only in the amount of explicit knowledge and it contributes to development in accuracy, even though the learners depend on both explicit knowledge and implicit knowledge when writing. Williams (2012) also stated that written CF affects development in explicit knowledge, not in implicit knowledge. There are no empirical studies directly tackling this issue in the field of 'written' CF studies, but there are some empirical studies treating the effects of 'oral' CF on development in implicit and explicit knowledge. Li et al. (2016) examined the effects of two types of 'oral' CF on development in both types of knowledge for the English past passive construction. One strategy was *corrective recasts* (Doughty & Varela, 1998), where erroneous utterances were immediately repeated with the error highlighted through emphasis to encourage self-correction, followed by recasts that reformulated the wrong utterance without altering the meaning. The other strategy was delayed feedback, which was provided to every error one by one after the completion of the task by the teacher, and which encouraged a learner to self-correct such as "Can you say it correctly?" When failed to self-correct, the learners were provided with a corrected linguistic form from the teacher. The results showed that both types of oral CF only improved the scores of the untimed grammaticality judgment test (untimed GJT), in other words, explicit knowledge. However, giving recasts led to improvement in accuracy in relatively free communication where implicit knowledge was demanded. Ellis (2004) explains that

immediate and delayed judgements in a GJT reflect implicit and explicit knowledge respectively. According to his detailed explanation, a GJT potentially involves three processing operations: *semantic processing*, *noticing*, and *reflecting*. In the stage of semantic processing, learners understand the meaning of a sentence. In the stage of noticing, they search to establish whether something is formally incorrect in the sentence, and in the last stage, reflecting, they consider what is incorrect about the sentence and, possibly, why it is incorrect. In a timed GJT, learners are allowed semantic processing and noticing, while an untimed GJT allows opportunity for all three processing operations semantic processing, noticing, and reflecting to take place. In addition, Gutierrez (2013) stated that learners' responses to grammatical and ungrammatical items load on separate factors, with the former tapping implicit knowledge and the latter explicit knowledge in addition to the existence of time pressure. Although a timed GJT should keep the participants on the access only to semantic processing and noticing, those who can quickly process are considered to further access to reflecting and to use explicit knowledge to identify what is incorrect and why it is incorrect. For this reason, only the sentences with no error were focused and the other ungrammatical sentences were not given attention. On the other hand, there is the possibility of using only implicit knowledge when learners judge grammatical sentences as grammatical, therefore, only the sentences including some errors were focused on for examining development in explicit knowledge. In addition, Meta-analysis of oral CF studies (Li, 2010; Lyster & Saito, 2010) indicated that some studies claimed that oral CF brought development in implicit knowledge though the effect size was small.

These findings, however, should be treated with care. Li et al. (2016) introduced a theory in cognitive psychology regarding delayed CF as stimulus to acquire implicit knowledge in order to explain the potential for direct contribution of written CF to development in implicit knowledge. According to reactivation and reconsolidation theory (Nader, 2003), when memories are reactivated in conditions that make them susceptible to change, their labile state allows for reconsolidation. This reconsolidation occurs not only in declarative or explicit but also in procedural or implicit memory knowledge. For example, when the linguistic explicit knowledge of some rules is activated when recalled, and corrected if the knowledge is inaccurate, accurate knowledge written CF offers is reconstructed or stored. This reconstruction is said to happen both in declarative or explicit knowledge and in procedural or implicit knowledge. Thus, this theory implies that procedural or implicit knowledge is directly acquired through written CF without repeated practice. The theory is

not established in the field of SLA, and there have not been enough empirical studies to illustrate that procedural knowledge *related to language* is really acquired by only reactivation and reconsolidation without practice *related to language*.

Moreover, the influence of Transfer-Appropriate Processing (TAP) should be also taken into consideration. According to Lightbown (2008), TAP theory claims that what we have learned can be best retrieved when the condition for retrieval matches the condition in learning. That is, there were more or less possibilities of influence of TAP in the studies mentioned above because the condition of treatment where recasts were given during communication matched the condition of tests which investigated the effects of recasts on implicit knowledge.

In short, we are lacking in empirical studies examining the effectiveness of written CF on development in explicit knowledge and implicit knowledge within a single research design, using appropriate measuring tools developed for examining two types of knowledge.

2.4.1.4 A Single Provision or Multiple Provisions

Kang and Han (2015) claimed that even a single treatment of written CF is effective for improving accuracy in the writing of new texts. Most empirical studies have focused on a single treatment and treated a narrow range of linguistic categories so far, so it is difficult to draw any conclusion about whether a single-shot written CF truly contributes to L2 development. In addition, we do not know much about how learner's knowledge and performance gradually change as they are given some opportunities to receive written CF through multiple treatments. In educational settings, it is natural for language teachers to offer written CF again and again on errors of the same linguistic categories. Making clear how multi-shot written CF strategies affect learner's L2 development and how the effectiveness of written CF gradually changes will be helpful for teachers.

2.4.2 Problem of Which Types of Written CF Are Compared

As stated earlier, the comparison of the effects of written CF has been mainly conducted between direct written CF and indirect written CF, and therefore have not fully included metalinguistic written CF.

In addition, the relative effectiveness of written CF has been studied on the basis of the dichotomy: either focused or unfocused. Many studies have targeted only one, two or three error

categories at one time, and found that focused written CF strategies facilitated accuracy. On the other hand, very little research has investigated the effectiveness of unfocused written CF, and the findings are contradictory (Truscott & Hsu, 2008; VanBeuningen et al., 2012). Irrespective of whether a single-shot focused written CF or a single-shot unfocused written CF is effective in learner's improvement in linguistic accuracy, the question of whether one of these approaches is more effective than the other can be answered only if the two strategies are compared within a single research.

Ellis et al. (2008) compared the effectiveness of focused and unfocused written CF on Japanese intermediate EFL learners in a single study. While the focused group received direct written CF only on the errors in the article, the unfocused group received direct written CF on the errors not only in the article but in other error categories. Although the researchers concluded that both types of feedback were equally effective, they acknowledged that they were not able to sufficiently distinguish one from the other because article errors appeared with high frequency in both. Sheen, Wright, and Moldawa (2009) investigated the effects of focused and unfocused written CF on article errors and on a broader range of grammatical structures (articles, copula 'be,' regular past tense, irregular past tense, prepositions). They reported that focused written CF on a single grammatical target (the English article system) alone was more effective than unfocused written CF, but at the same time, they admitted that the written CF given to the unfocused group was not so systematic; some of the errors were corrected but others were not.

Because of the limitations in both studies above, it is not possible to draw any conclusion about the superiority of focused written CF for or L2 development over unfocused written CF. Theoretically, it may be argued that learners with a more developed knowledge of the forms or structures may benefit from unfocused written CF, while learners with only partially developed knowledge may need more focused feedback if cognitive load is considered to be critical for L2 development. Thus, more empirical studies are needed.

2.4.3 Problem of Whether Individual Learner-Internal Factors Are Concerned

According to Sheen (2007), learners can vary enormously with regard to cognitive factors such as aptitude, intelligence, and proficiency, as well as affective factors such as language anxiety, motivation, and attitude. It has become clear that the effectiveness of written CF is mediated by such

individual factors. Therefore, it is important to consider a wide range of factors that might facilitate or impede the learner's cognitive processing of input. Factors that may impact upon cognitive processing include individual learner-internal cognitive factors (e.g., working memory and processing capacity), individual learner-internal motivational or affective factors (e.g., interest, attitudes, beliefs) and individual learner-external factors (e.g., pedagogical and instructional factors, social relationships). In this dissertation, individual learner-internal cognitive and affective factors are given focus, which are considered to have a strong influence on progress in cognitive processing for L2 development.

The individual cognitive factor of focus in this dissertation is L2 proficiency. It is the ability of an individual to speak or comprehend in the target language. It is largely related to the size of the learner's long-term memory storage including both explicit and implicit L2 knowledge, and to the working memory capacity related to language comprehension and production. Working memory is the site where new input is stored and incorporated with information already stored in long-term memory, which is said to be important in the processes such as attention, noticing, hypothesizing and restructuring. Unlike long-term memory, working memory has a limited capacity, and therefore is constrained by the amount of cognitive load in processing at one time. According to Skehan (1998), learners with larger working memory capacities are better equipped to attend to and process input, and prepare for output. In particular, lower proficiency learners may have great difficulty in attending to more than one aspect of language simultaneously. Because learners with a lower level of proficiency need to process new information in a more controlled manner, more effort and attention are needed in their working memory. In all stages of the information processing and production, it is expected that individual differences in L2 proficiency influence more or less any process in the stages. For example, the L2 proficiency level may determine if the CF is comprehended. If the learner has only partially stored information about when and why the linguistic form or structure in his or her long-term memory, explicit metalinguistic information may be most helpful to comprehend more clearly and fully. In the process of hypothesis testing, it may be that the working memory has a less crucial role to play in the processing of written CF where learners can refer to what they wrote and to what the feedback says many times, and obtain the greater amount of time for analysis than in that of oral CF where the engagement period is fleeting (Williams, 2012). One advantage of hypothesis-testing that results from processing 'written' CF

may be that learners feel more comfortable doing it than they feel in the settings where issues of face and identity may be more threatened if hypotheses prove to be incorrect, which would usually happen in oral communication. There are many factors that can explain why a learner fails to produce an accurate output on certain occasions. When they produce the target language, they need to have attentional control over the production of meaning and appropriate form and structure, and to retrieve the newly integrated knowledge from the long-term memory. This processing in output requires the working memory capacity, and therefore it is influenced largely by L2 proficiency.

Studies 1 to 4 in this dissertation try to identify the most effective written CF according to the learner's language proficiency level, one of the individual learner-internal factors. In Study 5, the learner's attitude toward corrective feedback strategies and text revisions is focused on, which is one of the individual learner-internal affective factors considered to influence their receptivity to error correction, and thus the effectiveness of the feedback.

Although attitudes to language learning in general, to target language communities, and to learning of a particular target language have also been identified in the SLA literature as affective factors, little attention has been given to the way L2 learners respond to written CF and text revisions. They might affect whether or not learners are ready and willing to attend to accuracy and to written CF, and engage in cognitive processing activities such as noticing the gap and hypothesis testing. For example, if they have prior experiences that written CF did not enable learners to accurately modify linguistic errors, they may decide to ignore written CF. Motivated learning behavior would seem to be necessary for learners to consolidate their renewed knowledge so that it can be retrieved automatically from their long-term memory over time.

CF research into learners' attitudes has been mainly descriptive so far, identifying learners' perceptions and preferences to certain types of feedback. Leki (1991) studies ESL students' preferences for error correction and found that they wanted to write errorless English and considered their teacher as the best source of error correction. Regarding the students' preferences for the type of CF, about seventy percent of the students asked for indirect CF which indicates the location of the error together with metalinguistic clues to help them to correct the error by themselves. Twenty-five percent of the students considered direct CF providing the corrected error as most desirable. Lastly, no students approved of indirect CF.

Enginarlar (1993) investigated students' feelings about the utility and instructional value of

written CF. He found that most students highly valued the teacher's CF on their written compositions, which agrees with Leki's (1991) finding, while the students did not favor revision exercises. In addition, Schulz (1996, 2001) reported that ESL students and FL students viewed grammar instruction and corrective feedback as very important for learning a second or foreign language. However, it is not clear whether what learners prefer and desire is actually what is best for language development. Thus, what is needed is empirical studies that examine the relationship between learners' attitudes toward error correction and text revisions, and actual language learning resulting from CF.

2.4.4 Other Problems and Limitations in Research Design
2.4.4.1 Linguistic Category Treated

It is also said that targeted linguistic categories have been very limited and almost all of the studies have dealt with English article systems. Conditionals (Shintani et al., 2014) and preposition (Guo, 2015) have been focused, but more research that deals with a wide range of linguistic categories is asked for, which will provide useful information for language teachers.

2.4.4.2 Scientific Method Used

Empirical and scientific research, which asks researchers to plan, conduct and analyze the study adequately by, for example, controlling various factors and adopting the pre-post-delayed-posttest design, is also needed. This is partly because the research on written CF has been conducted mainly within the pedagogical domain of L2 writing. L2 writing research has focused on feedback to the contents of a written text as well as to errors in linguistic forms, while SLA has only focused on the linguistic errors. Interest for written CF from SLA researchers has emerged relatively recently. L2 writing research is mainly interested in how written CF contributes to development in learners' editing strategies in writing, that is, development in self-correction of the first draft, where SLA pays more attention to the linguistic development, that is, development in accuracy in new pieces of writing. Therefore, it is not clear whether written CF really contributes to L2 development, and more studies adopting a pre-post-delayed-posttest research design and a control group are needed.

To sum up, problems and limitations in the previous studies are as follows:

(1) Many studies have examined the effects of written CF on new pieces of writing and on text revisions separately.

(2) The effects of written CF have not been fully examined on the basis of different types of new texts.

(3) Many studies have focused on the effects of written CF on development in explicit knowledge, not in implicit knowledge.

(4) The effects and gradual changes of effects caused by multiple provisions of written CF are not clear.

(5) The comparative studies on the effects of written CF have mainly treated the difference between direct and indirect written CF. Metalinguistic written CF has not been treated so frequently.

(6) There are few comparative studies conducted within one single research design on the effects of focused and unfocused written CF.

(7) The individual learner-internal cognitive or affective factors which would mediate the effectiveness of written CF, such as English proficiency and learner's attitudes toward written CF and text revisions, have not been fully included so far.

(8) The range of targeted linguistic categories is narrowly limited.

(9) The studies have sometimes ignored a control group or the pre-post-delayed-posttest research design for examining the effects of written CF on L2 development.

2.5 Aim of the Dissertation

The overarching aim of this dissertation is to identify the most effective written CF according to learners' levels of L2 proficiency. In order to accomplish this purpose, the relative effectiveness of written CF is examined under different circumstances from theoretical and pedagogical perspectives. More specifically, the purpose is to compare and clarify the effects of written CF strategies on improvement in text revisions and the writing of new texts, on improvement in different types of tests, and on development in both explicit and implicit knowledge, dividing the proficiency into mainly two levels, higher or lower, which could be one of the mediating factors influencing the effectiveness. Additionally, the study also focuses on learners' attitudes toward

written CF and text revisions, which are also one of the mediating factors, in order to consider the relationship between the effectiveness of written CF and their affective states. In order to accomplish these purposes, five individual studies are designed and conducted, which will be explained in Chapters 3 to 7 respectively in more detail.

2.6 Structure and Focus of the Dissertation

The first study (Study 1) reported in Chapter 3 examines the effectiveness of two types of written CF, i.e., direct written CF and metalinguistic written CF, on development in text revisions and on new pieces of writing, which have been often treated separately, in a single research containing a control group. In addition, the study manifests how test scores and the ratio of successful self-correction change, given multiple episodes of providing written CF. Therefore, L2 development through written CF is measured on the basis of an increase in accuracy on the tests for new pieces of writing, and of an increase in the ratio of successful self-correction for text revisions. The targeted grammatical categories are the conditionals including the future conditional, the present-counterfactual conditional, the past-counterfactual conditional, and the mixed-counterfactual conditional, and one type of test, an English translation test (ETT), is adopted. The learners are divided into two groups according to their English proficiency levels, higher or lower. The learners at each level of proficiency are further divided into three groups, i.e., the metalinguistic written CF group, the direct written CF group, and the control group. In order to examine the effects as an editing tool of written CF on text revisions, the ratio of successful self-correction is calculated, and in order to examine the effects as a learning tool on a new piece of writing, the test scores are calculated.

The second study (Study 2) reported in Chapter 4 investigates the relative effectiveness of written CF on three kinds of tests. One test examines the effects on the acquisition of accurate grammatical knowledge, and others in performance in new writing tasks. In this study, the effects on the acquisition of accurate grammatical knowledge can be equivalent to the effects on the acquisition of explicit knowledge caused by reassessment and hypothesis reforming, which are also investigated in Study 3. The effects of written CF on new performance are investigated not with a

single task, but with two tasks demanding different amount of the working memory capacity. L2 development is measured on the basis of an increase in accuracy on the tests. The targeted grammatical categories are the conditionals including the future conditional, the present-counterfactual conditional, and the past-counterfactual conditional. Three different measuring tools are adopted: the untimed GJT for measuring the acquisition of accurate explicit knowledge, and the ETT and the essay writing test (EWT) for measuring the improvement in accuracy in the writing of new texts. The learners are divided into two groups according to their levels of English proficiency, higher or lower. The learners in each proficiency group are further divided into three groups: the focused metalinguistic written CF group, the focused direct written CF group, and the unfocused direct written CF group.

The third study (Study 3) reported in Chapter 5 measures the relative effectiveness of different written CF, direct CF and metalinguistic CF, on the basis of the direct contribution to development in explicit knowledge and that in implicit knowledge from the perspective of SLA. The effectiveness of written CF is examined, depending on two different levels of English proficiency, higher and lower. The targeted grammatical category is the present perfect tense, whose meaning and structure in the sentence are assumed to be difficult to understand and produce, confused with those of the past tense. Two different types of tests for measuring development in implicit knowledge; the timed GJT, where learners have to judge the grammaticality of each sentence quickly, and the elicited imitation test (EIT), where learners have to reproduce the sentence they listen to and where they cannot enjoy the benefits of the influence of TAP, are developed. The test for measuring development in explicit knowledge is the untimed GJT. The former test includes three groups, the metalinguistic written CF group, the direct written CF group, and the control group, and follows the pre-post-delayed-posttest research design, while the latter includes only two groups; the metalinguistic written CF group and the control group only with the pre-posttest design.

The fourth study (Study 4) in Chapter 6 examines the relative effectiveness of different types of written CF on an increase in accuracy in new pieces of writing, taking the learner's grammatical item-specific proficiency into consideration, which is originally named and defined for this study. As explained so far, in this dissertation, proficiency means the size of learner's long-term memory store and working memory capacity that relate to both comprehension and production of the target language. Thus, strictly speaking, we can propose that each learner has a different level of

Table 2.3

Focus of Studies 1 to 5

Focus	Study 1	2	3	4	5
1. Development in revisions and new writing	✓				
2. Development in different tests		✓			
3. Development of explicit and implicit knowledge			✓		
4. Single treatment and multiple treatments	✓				
5. Metalinguistic CF	✓	✓	✓	✓	
6. Focused CF and unfocused CF		✓			
7. Learner-internal factors	✓	✓	✓	✓	✓
8. Different grammatical category	✓	✓	✓	✓	
9. Scientific method	✓	✓	✓	✓	

proficiency in each grammatical category. The difference between higher and lower proficiency levels in Studies 1 to 3 depends on general L2 proficiency, which is decided by the scores of three or four skill-based English tests, while this Study 4 adopts proficiency determined by the test scores according to each grammatical category. The focused grammatical categories are the present perfect tense and the past perfect tense, and the ETT is adopted as a measuring tool. The participants are divided into mainly three groups according to the test scores: a higher item-specific proficiency group selected by the results of the writing test for the present perfect tense, a middle item-specific proficiency group and a lower item-specific proficiency group, both of which were selected by the results of the writing test for the past perfect tense. Higher or middle item-specific proficiency group is further randomly divided into four groups, the direct written CF group, the indirect written CF group, the metalinguistic written CF group, and the control group respectively, while a lower item-specific proficiency group into three groups, the direct written CF group, the metalinguistic written CF group, and the control group (no indirect written CF group).

The last study (Study 5) reported in Chapter 7 focuses on an affective mediating factor, i.e.,

learner's attitude toward written CF and text revisions, and examines how responses in the questionnaire differ according to their levels of proficiency. Four questions are developed: (1) *Who do you want to correct your errors?*; (2) *How do you want your errors to be corrected?*; (3) *How many errors do you want to be corrected?*; and (4) *What do you do after receiving written CF?* By referring to the results, the relationship between the effectiveness of written CF, which turned out to be clear through Studies 1 to 4, and learners' affective attitudes, which manifested in Study 5, is considered. Focuses of each study are summarized in Table 2.3. The participants in Studies 1 to 5 were requested the cooperation in them in advance. Studies 1 to 5 were conducted with their permission.

In Chapter 8, the main findings gained through Studies 1 to 5 are summarized first, which is followed by a discussion of what the five studies clarify as to the contribution of written CF to L2 development. Then, the pedagogical implications are stated, which will be useful for classroom teachers looking for its value in practical use. Finally, this dissertation will close with the introduction of problems and limitations found in Studies 1 to 5, and with some recommendations for further research.

CHAPTER 3

Study 1: The Effectiveness of Written CF on Text Revisions and the Writing of New Texts

As explained in Chapter 2, there is a difference in information that each written CF offers. Direct written CF offers the information about a correct linguistic form for each error, which is thought to be helpful in text revisions because learners have a chance to directly use the form, even if they do not understand the rule. However, in the case of the writing of new texts, direct written CF may not be helpful because they have to understand how the form are used. On the other hand, metalinguistic written CF, which provides metalinguistic information not only forms but also rules, would be helpful in a new piece of writing. For this reason, the relative effectiveness of written CF must be examined bot in text revisions and in new pieces of writing. In addition, most empirical studies have focused on a single treatment of the provision of written CF, and treated a narrow range of linguistic categories so far, so it is difficult to draw any conclusion about whether a single treatment of written CF truly contributes to L2 development. Furthermore, we are lacking in empirical studies that investigate how multiple treatments of written CF influence L2 development and how the effectiveness of written CF gradually changes.

3.1 Research Questions

Four research questions (RQs) were addressed to investigate the relative effectiveness of two types of feedback (metalinguistic written CF and direct written CF) on text revisions and new pieces of writing according to learners' levels of proficiency (higher and lower) within a single research design. At the same time, this study tried to clarify how the effectiveness of written CF changes through multiple provisions of written CF. The grammatical categories of focus were four types of the conditionals (the future or predictive conditional, the present-counterfactual conditional, the past-counterfactual conditional, and the mixed-counterfactual conditional). The measuring tool was

an English translation test (ETT):

RQ 1: Does written CF lead to the improvement in accuracy in text revisions?

RQ 2: If so, is there any difference in the effects of written CF depending on English proficiency level?

RQ 3: Does written CF lead to the improvement in accuracy in new pieces of writing?

RQ 4: If so, is there any difference in the effects of written CF depending on English proficiency level?

3.2 Method

3.2.1 Participants

A total of 110 Japanese learners of English in high school participated in this study. They were all second-year high school students and had received at least 5 years of formal English instruction at their junior and high schools. When first-year students, they were supposed to choose their learning course of English, standard or advanced course, on the basis of each individual's free will, and they were not allowed to change their course. In this study, 52 learners in an advanced course are considered as being in the higher English proficiency group, and 58 learners in a standard course as being in the lower English proficiency group. They all took an advanced version of the English test called *GTEC for STUDENTS* by Benesse Corporation, whose maximum score is 810, before participating in this study. The means in total score were 680.5 ($SD = 48.25$) for the higher proficiency group and 496.2 ($SD = 16.71$) for the lower proficiency group. The difference between them in the means was statistically significant ($F (1,108) = 732.18$, $p < .01$). Considering only the scores in writing whose maximum value is 170, the scores the learners in the higher English proficiency group got ($M = 132.9$, $SD = 10.11$) were significantly higher ($F (1,108) = 39.09$, $p < .01$) than those in the lower English proficiency group ($M = 121.7$, $SD = 8.40$). In each proficiency level, the learners were divided into three groups; in the case of the higher proficiency level, the learners were assigned to the metalinguistic written CF group ($n = 15$), the direct written CF group ($n = 16$), and the control group ($n = 21$). In the same way, lower proficiency learners were appointed to the

metalinguistic written CF group ($n = 19$), the direct written CF group ($n = 23$), and the control group ($n = 16$). Indirect written CF, which was generally proved to have a smaller effect than direct written CF, was not be included in this study.

3.2.2 Target Structures

The target structures in this study were four types of the conditionals; the future conditional, the present-counterfactual conditional, the past-counterfactual conditional, and the mixed-counterfactual conditional. The future conditional is mainly used to express future plans or outcome, whose normal pattern is simple present tense in the *if*-clause and some explicit indication of future time in the main clause. Counterfactual conditionals refer to impossibilities with reference to the present or the past. The present-counterfactual conditional consists of simple past tense or present subjunctive in the *if*-clause and *would* in the main clause, while the past-counterfactual consists of the past perfect tense in the *if*-clause and *would* be followed by perfect aspect. The present- and past-counterfactual is a mixed version, and consists of *would* in the main clause, and the past perfect tense in the *if*-clause. According to Larsen-Freeman and Celce-Murcia (2016), conditional sentences consist of two clauses, and therefore are more syntactically complex than other structures. Furthermore, the semantics of all types of conditionals is subtle and hard to understand especially for L2 learners. Even for higher English proficiency learners, the structures are difficult to comprehend and produce accurately, which means they impose heavy cognitive load on the learners, and which also means the learners are likely to make errors in writing. Examples of four types of the conditionals used in this study are as follows:

(1) If it rains tomorrow, we will stay home.　　　　(future conditional)
(2) If he were free, he would help you.　　　　(present-counterfactual conditional)
(3) If she had had ten million yen, she would have bought a yacht. (past-counterfactual conditional)
(4) If I had worked harder, I would be happier now.　　(mixed-counterfactual conditional)

3.2.3 Design

During Week 1, the participants completed the ETT as the pretest (Pretest) after taking a 90-minute English lesson where they received an explicit explanation of the target structures and did

some drills for checking comprehension of the structures. In Week 2, each group, i.e., the metalinguistic written CF, the direct written CF, and the control groups, had a chance to revise the first test and performed the second ETT. Then, they received written CF with an opportunity to revise the second test except for the control group. In Week 3, each group completed the same kind of test (the third ETT), received written CF again, and revised the third test. That is, a revised handout for Pretest became Revision 1, and a revised handout for the first posttest (Posttest 1) became Revision 2. Finally, a revised handout for the second posttest (Posttest 2) became Revision 3.

3.2.4 Testing and Treatment Materials

Pretest, Posttest 1, and Posttest 2 were the ETTs (Appendix A). Each test consisted of twelve questions where the participants had to use three future conditionals, three present-counterfactual conditionals, three past-counterfactual conditionals, and three mixed conditionals in order to complete the writing test. Around fifteen minutes were assigned to the test for every learner to fully refer to their linguistic knowledge and to give a second look. In order to keep a balance of difficulty among three tests, only vocabulary was changed with keeping the sentence structures intact. In case the participants were not able to find a base form for each verb to complete a sentence, verbs and other English vocabulary which seemed to be difficult for the participants to recall were put on the section named *Words* on the handout of the test in advance. Scoring was conducted on the main clause and the *if*-clause separately. For example, in the case of a sentence required for the use of the present-counterfactual conditional, whether the word *if* and simple past tense are correctly used in the *if*-clause, and whether the past tense in the auxiliary verb and a base form of a verb are precisely used in the main clause were thoroughly examined. One point was given to each errorless clause, while no point and only the mark of X to each incorrect clause. Thus, the maximum score was twenty-four points (two points for each sentence). Errors on which the study does not focus, such as those in spelling, the article, or the plural form of nouns, were excluded from the targets of scoring.

When there was an error, different kind of written CF was given, according to the group the learners belonged to. In the direct written CF group, the learners received a handout listing all of the correct forms. In the metalinguistic written CF group, either a circle ('correct') or an X mark ('incorrect') was given to the main clause and the *if*-clause respectively, and if 'incorrect', the sign

like *Check 1* was put around the X mark. Furthermore, in this group, the handout named a feedback sheet, whose size was A4, was distributed to the learners (Appendix B). With the sheet, the learners can find information about differences among the conditionals in addition to linguistic rules for each correct usage along with some examples. They cannot, however, find information about a correct form itself (an answer) to each question. Providing CF in the form of a 'sheet' would save time in classroom settings, while it would have a risk that a learner cannot find information necessary for him or her to notice the gap and to self-correct on the sheet. That is, there is a danger for the information not to be attended to by the learners and not to function as 'corrective' feedback to foster noticing the gap, especially for the lower English proficiency learners, who tend to make enormous errors at one time. To avoid this, the sign like '*Check 1*' was placed near each error as stated above. The number on the sign written in the worksheet was linked to the number described on the feedback sheet. For instance, when a learner receives the sign '*Check 1*' on a worksheet, she or he can refer to the information labeled '*Check 1*' which gives useful scaffolded help for self-correction. Each participant was asked to consider each error, comparing it with the information on written CF. After Posttest 2 and Revision 3, every participant in each group took a 50-minute English lesson to take advantage of an equal opportunity of learning, where the feedback sheet for the direct written CF group and the answer sheet for the metalinguistic written CF group and both sheets for the control group were offered.

3.2.5 Data Analysis

The scores on Pretest, Posttest 1, and Posttest 2 were subjected to a series of statistical analyses for the analysis of the effects of written CF on the writing of new texts. In addition, the ratios of successful self-correction through Revisions 1, 2 and 3 were subjected to a series of statistical analyses for the analysis of the effects of written CF on text revisions. If a learner made ten errors in total and succeeded in correcting five errors with written CF, the ratio of self-correction gained by dividing the number of successful self-correction by the total number of errors, 0.5, was given to the learner as a score. A repeated-measures ANOVA analyzed the comparative effects of the treatment for each test score and each ratio. One-way ANOVA with Holm's post hoc pair-wise comparisons was used to isolate the exact points in time where differences between the groups occurred when there was a significant Time x Group effect. Effect sizes for the ANOVA were

estimated as partial eta-squared ($\eta_p{}^2$). Effect sizes for the pairwise comparisons were estimated using Cohen's *d* with values of .20, .50, and .80 indicating small, medium, and large effects, respectively (Cohen, 1988).

3.3 Results

This section first reports the comparative effects of written CF on text revisions according to the levels of English proficiency (RQs 1 and 2). Then, it reports the comparative effects of written CF on the writing of new texts according to the levels of English proficiency (RQs 3 and 4). All tables of ANOVA in this study are shown in Appendix C.

3.3.1 Effects of Written CF on Text Revisions
3.3.1.1 Higher English Proficiency Group

Table 3.1 shows the descriptive statistics for scores for the two treatment groups (the metalinguistic written CF and the direct written CF groups) at the revisions of Pretest (Revision 1), the posttest (Revision 2), and the delayed posttest (Revision 3) in the higher English proficiency group. A repeated-measures ANOVA showed that there was a statistically significant effect only for Group ($F_{(2, 49)} = 57.36$, $p < .01$, $\eta_p^2 = .701$), while there were no significant effects for Time (F (2,

Table 3.1

Descriptive Statistics for Revisions 1 to 3 (Higher Proficiency Group)

Groups	*n*	Revision 1 Mean (*SD*)	Revision 2 Mean (*SD*)	Revision 3 Mean (*SD*)
MCF	15	0.81 (0.35)	0.90 (0.22)	0.97 (0.06)
DCF	16	0.95 (0.13)	0.89 (0.25)	0.92 (0.25)
*NF (Control)	21	0.31 (0.38)	0.16 (0.29)	0.15 (0.35)

Note. NF = No Feedback
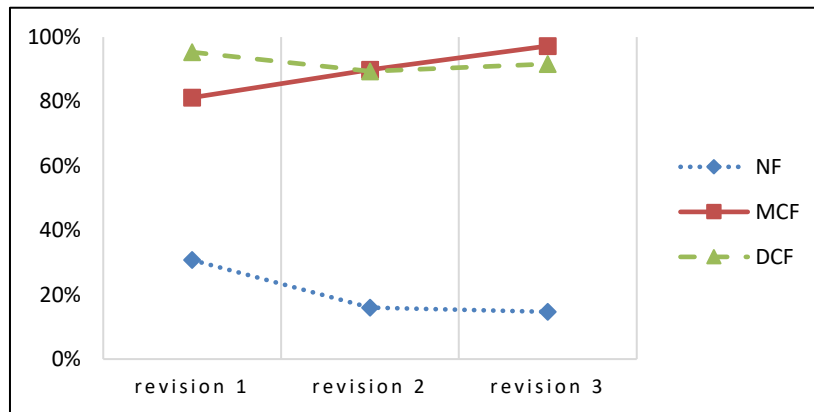
Figure 3.1. Group means of the ratio of successful self-correction among higher proficiency learners.

98) = 0.42, *ns*, η_p^2 = .008), and for Time x Group interaction (*F* (4, 98) = 2.28, *p* < .10, η_p^2 = .085) (Figure 3.1). It is said from this result that written CF had positive effects for higher proficiency learners on text revisions, but the difference between the metalinguistic written CF group and the direct written CF group was not clearly identified because of a ceiling effect.

3.3.1.2 Lower English Proficiency Group

Table 3.2 shows the descriptive statistics for scores for the two treatment groups (the metalinguistic written CF and the direct written CF groups) at the revisions of Pretest (Revision 1), the posttest (Revision 2), and the delayed posttest (Revision 3) in the lower English proficiency

Table 3.2

Descriptive Statistics for Revisions 1 to 3 (Lower Proficiency Group)

Groups	*n*	Revision 1 Mean (*SD*)	Revision 2 Mean (*SD*)	Revision 3 Mean (*SD*)
MCF	19	0.55 (0.24)	0.66 (0.37)	0.81 (0.24)
DCF	23	0.95 (0.14)	0.94 (0.12)	0.86 (0.26)
NF (Control)	16	0.05 (0.07)	0.00 (0.00)	0.03 (0.11)
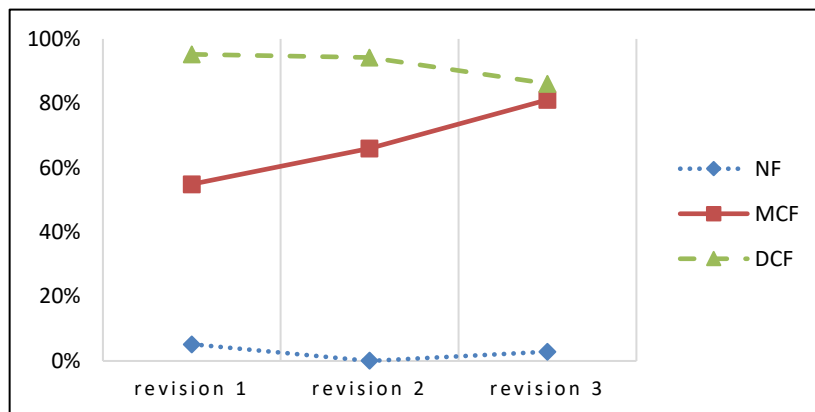
Figure 3.2. Group means of the ratio of successful self-correction among lower proficiency learners.

group. A repeated-measures ANOVA showed that there was no statistically significant effect for Time (F (2, 110) = 1.47, *ns*, η_p^2 = .026). However, there were statistically significant effects both for Group (F (2, 55) = 156.39, $p < .01$, η_p^2 = .850) and for Time x Group interaction (F (4, 110) = 7.20, $p < .01$, η_p^2 = .207). Holm pairwise comparisons showed that the significant group differences were found in Revisions 1 to 3. In Revision 1, the direct written CF group showed a significant advantage over the metalinguistic written CF group with a large effect size (d = 2.09) and over the control group with a large effect size (d = 7.71). In Revision 2 as well, the direct written CF group showed a significant advantage over the metalinguistic written CF group with a large effect size (d = 1.06) and over the control group with a large effect size (d = 10.16). In Revision 3, however, the significant difference between the metalinguistic written CF group and the direct written CF group was not found, which means the provision of metalinguistic written CF three times improved the ratio of self-correction to the same extent as direct written CF in the case of learners with a lower English proficiency (Figure 3.2).

3.3.2 Effects of Written CF on the Writing of New Texts

Table 3.3

Descriptive Statistics for the Test (Higher Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest 1 Mean (*SD*)	Posttest 2 Mean (*SD*)
MCF	15	21.00 (3.79)	20.80 (3.97)	21.93 (2.82)
DCF	16	17.81 (5.38)	18.94 (4.28)	19.69 (3.70)
NF (Control)	21	17.00 (5.15)	17.19 (4.85)	15.14 (5.12)

3.3.2.1 Higher English Proficiency group

Table 3.3 shows the descriptive statistics for test scores for the two treatment groups (the metalinguistic written CF and the direct written CF groups) at the three tests (Pretest, Posttest 1, and Posttest 2) in the higher English proficiency group. A repeated-measures ANOVA showed that there was no statistically significant effect for Time (F (2, 98) = 0.38, *ns*, η_p^2 = .008). However, there were statistically significant effects both for Group (F (2, 49) = 5.48, $p < .01$, η_p^2 = .182) and for Time x Group interaction (F (4, 98) = 3.80, $p < .01$, η_p^2 = .134). Holm pairwise comparisons showed that
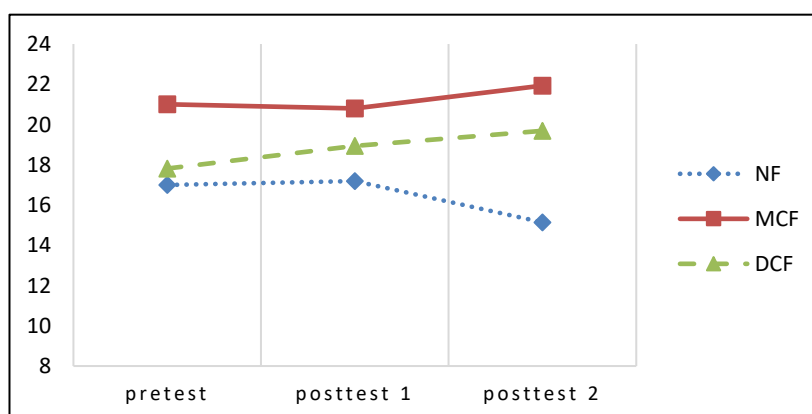
Figure 3.3. Group means of the conditionals among higher proficiency learners.

Table 3.4

Descriptive Statistics for the Test (Lower Proficiency Group)

		Pretest	Posttest 1	Posttest 2
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
MCF	19	11.84 (3.70)	15.47 (5.66)	16.21 (5.34)
DCF	23	11.04 (5.47)	11.96 (5.89)	12.09 (4.09)
NF (Control)	16	9.75 (4.66)	10.94 (4.29)	10.00 (4.74)

the significant group differences were not found in Pretest and Posttest 1. In Posttest 2, the metalinguistic written CF group showed a significant advantage over the control group with a large effect size ($d = 1.57$) and the direct written CF group also had a significant advantage over the control group with a large effect size ($d = 1.00$). There was no significant difference between the metalinguistic written CF group and the direct written CF group. However, the significant difference between the two experimental groups (the metalinguistic written CF and the direct written CF groups) and the control group could be caused by the decrease of scores in the control group. Thus, it can be safe to say that the differences of the effects of written CF on the improvement in accuracy in new pieces of writing in the higher English proficiency group were not recognized (Figure 3.3).

3.3.2.2 Lower English Proficiency Group

Table 3.4 shows the descriptive statistics for test scores for the two treatment groups (the metalinguistic written CF and the direct written CF groups) at the three tests (Pretest, Posttest 1, and Posttest 2) in the lower English proficiency group. A repeated-measures ANOVA showed that there were statistically significant effects for Time ($F_{(2, 110)} = 9.62$, $p < .01$, $\eta_p^2 = .149$), for Group ($F_{(2, 55)} = 4.22$, $p < .05$, $\eta_p^2 = .133$) and for Time x Group interaction ($F_{(4, 110)} = 3.44$, $p < .05$, $\eta_p^2 = .111$). Holm pairwise comparisons showed that the significant group differences were not observed in Pretest and Posttest 1. However, in Posttest 2, the metalinguistic written CF group showed a significant advantage over the direct written CF group with a large effect size ($d = .88$) as well as over the control group with a large effect size ($d = 1.22$).
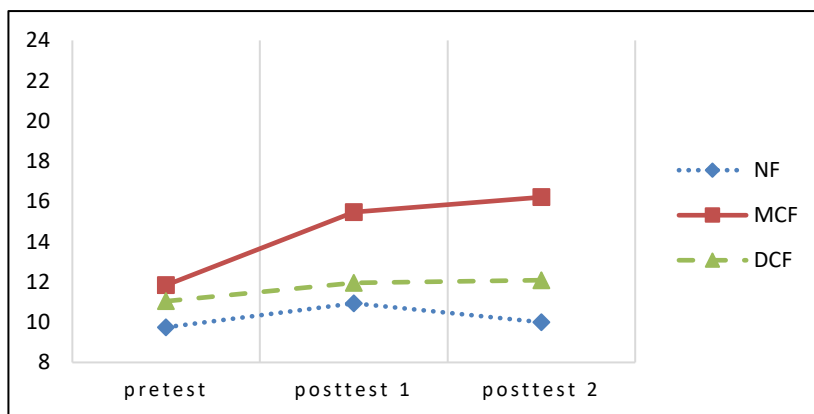
Figure 3.4. Group means of the conditionals among lower proficiency learners.

To sum up, the metalinguistic written CF treatment proved to be the most effective when it is offered in multiple occasions for the lower English proficiency group (Figure 3.4).

3.4 Discussion

Learners who receive direct written CF obtain information not of accurate 'rules' but of accurate 'forms.' For this reason, they need to inductively find the rules with the help of the forms given by direct written CF, and it is not clear whether or not the learner succeeds in really finding them, and, even if he or she succeeds, it is not clear whether or not the linguistic information about forms and rules integrated in long-term memory are correct or acceptable. That is, learners potentially store a correct linguistic form for each error and renewed linguistic knowledge about forms and rules in the stage of integration.

On the other hand, in the case of metalinguistic written CF, they can obtain not a correct form, but metalinguistic information about forms and rules, and therefore they should deduce a correct form that fits to each occasion. However, it is no clear whether they can really do such a thing or whether the form they deduce is truly correct. It is possible for learners to store metalinguistic

information including a correct rule and renewed information about forms and rules, at the integration stage.

RQ 1 asked whether written CF affected accurate revisions of the conditionals, and RQ 2 asked whether there was any difference in the effects of written CF according to learners' English proficiency levels if the answer to RQ 1 was *yes*. These questions were answered by examining the ratio of successful self-correction in the ETT dividing the proficiency level into two; lower or higher. The answer to RQ 1 was *Yes*. However, it proved that any written CF seemed to have no clear relative effectiveness on their revised texts in the higher English proficiency learners. Although the metalinguistic written CF group and the direct written CF group had a significant advantage over the control group, it was hard to make any conclusion because of a ceiling effect. On the other hand, in the lower English proficiency group, the direct written CF group outperformed the metalinguistic written CF group in Revisions 1 and 2. However, in Revision 3, the metalinguistic written CF treatment led to the ratio of successful self-correction to the same extent as the direct written CF treatment.

As explained in the previous chapter, Corder (1967) made a distinction between *errors* and *mistakes*. The former represents errors that occurred as a result of a lack of knowledge, while the latter merely performance phenomena reflecting processing failure. When learners receive accurate forms through direct written CF, they can correct errors in the phase of revisions by themselves, recalling the forms, even though they do not understand the linguistic rules behind forms. For this reason, it is natural that the direct written CF treatment had a positive effect on text revisions from the very first trial, irrespective of which types of *errors* (errors or mistakes) they made, and of whether their proficiency level was higher or lower. Furthermore, for learners with a higher level of proficiency, not only direct written CF but also metalinguistic written CF had positive effects on an increase in accuracy in revisions. Because those who belonged to this proficiency level were originally thought to have already stored a significant number of explicit linguistic rules of the target structures, the conditionals, they were able to deduce or recall the rules from the forms that direct written CF gave, and then to self-correct. On the contrary, it is assumed that errors made by learners with a lower level of proficiency tended to be errors not mistakes, and that they did not store accurate linguistic rules and if any, the rules could be inaccurate even they had explicit instruction in advance. For these reasons, a single-shot metalinguistic written CF was not efficient. However, given multiple

opportunities to refer to the feedback sheet, they would understand the rules, find the reasons of errors, deduce accurate linguistic forms, and as a result improve the ratio of successful self-correction.

RQ 3 asked whether written CF affected accurate usage of the conditionals in new pieces of writing, and RQ 4 asked whether there was any difference in the effects of written CF according to learners' levels of English proficiency. The answer to RQ 3 was *yes*, but only in the lower English proficiency level. In the higher English proficiency level, the two experimental groups (the metalinguistic written CF and the direct written CF groups) did not show any significant advantage over the control group, even though they were given written CF several times, while in the lower English proficiency level, the metalinguistic written CF group showed a significant advantage over the direct written CF and the control groups in Posttest 2.

In order for the learners to gain high scores in new pieces of writing, in contrast to text revisions, they have to understand accurate linguistic rules with which they can then deduce accurate linguistic forms. Even when they were given direct written CF to errors in the conditionals and could store accurate forms, there was no chance to use the same forms in new writing later. As stated above, the learners with a higher level of proficiency were thought to already store some explicit knowledge of the conditionals. Accordingly, a significant difference among the groups was not observed in this proficiency group regardless of the type of written CF, or of the existence of written CF. On the other hand, the learners with a lower level of proficiency would not have stored so many linguistic rules, and would have had great difficulty in deducing the accurate rules by themselves with direct written CF. It is for this reason that metalinguistic written CF giving accurate linguistic rules directly was more effective than direct written CF in the proficiency group.

CHAPTER 4

Study 2: The Effectiveness of Focused and Unfocused Written CF Strategies on Different Tasks

In classroom, learners perform many kinds of writing tasks which differ in cognitive load on the working memory capacity, which is one of the main constructs of L2 proficiency. Because the working memory capacity is considered to affect the effectiveness of written CF, the effectiveness should not be investigated only in a single writing task. However, we are lacking the studies on the effectiveness of written CF on improvement in different kinds of tasks. In addition, many studies have targeted focused written CF, however, very little research has investigated the effectiveness of unfocused written CF, which a number of teachers tend to adopt in correcting learners' errors. Moreover, we are lacking in the empirical studies comparing the effectiveness of focused written CF and that of unfocused written CF for L2 development within a single research design.

4.1 Research Questions

Two RQs were addressed to investigate the relative effectiveness of three types of written CF (focused direct written CF, unfocused direct written CF and focused metalinguistic written CF) on an increase in accuracy in three types of the conditionals (the future or predictive conditional, the present-counterfactual conditional, and the past-counterfactual conditional) through three different tests (an untimed grammaticality judgment test (GJT), an English translation test (ETT), and an essay writing test (EWT)) according to learner's levels of proficiency (higher and lower) within a single research design:

RQ 1: Does written CF lead to an increase in accuracy in three different kinds of tests?
RQ 2: If so, is there any difference in the effects of written CF depending on English proficiency level?

4.2 Method

4.2.1 Participants

A total of 141 Japanese learners of English in high school participated in this study. They were all third-year high school students and had received at least 6 years of formal English instruction at their junior and high schools. When first-year students, they decide their learning course of English, standard or advanced. In this study, 63 learners in an advanced course are considered as being in the higher English proficiency group, and 78 learners in a standard course as being in the lower English proficiency group. They all took the advanced version of *GTEC for STUDENTS* by Benesse Corporation, which focuses on four skills, and whose maximum score is 1280, before participating in this study. The means in total score were 962.7 (SD = 94.65) for the higher proficiency group, and 814.6 (SD = 73.37) for the lower proficiency group. The difference in the means was statistically significant (F (1,139) = 107.98, p < .01). Considering only the scores in writing whose maximum value is 320, the scores the learners in the higher English proficiency group got (M = 243.3, SD = 18.80) was significantly higher (F (1,139) = 26.51, p < .01) than those in the lower English proficiency group (M = 222.6, SD = 26.89). In each proficiency level, the learners were divided into three groups; in the case of the higher proficiency level, the learners were assigned to the focused metalinguistic written CF group (n = 23), the focused direct written CF group (n = 21), and the unfocused direct written CF group (n = 19). In the same way, in the case of the lower proficiency level, the learners were appointed to the focused metalinguistic written CF group (n = 29), the focused direct written CF group (n = 27), and the unfocused direct written CF group (n = 22). Unfocused metalinguistic written CF, which can give metalinguistic information about rules or forms to *every* error which each learner makes, was not included in the study because it seemed to be difficult to be operationalized, and indirect written CF, which is said to generally have a smaller effect than direct written CF, was not also included.

4.2.2 Target Structures

The target structures in this study were the future conditional, the present-counterfactual

conditional, and the past-counterfactual conditional. The future conditional is mainly used to express future plans or outcome, whose normal pattern is simple present tense in the *if*-clause and some explicit indication of future time in the main clause. Counterfactual conditionals refer to impossibilities with reference to the present or the past. The present-counterfactual conditional consists of simple past tense or present subjunctive in the *if*-clause and *would* in the main clause, while the past-counterfactual consists of the past perfect tense in the *if*-clause and *would* be followed by perfect aspect. According to Larsen-Freeman and Celce-Murcia (2016), conditional sentences consist of two clauses, and therefore are more syntactically complex than other structures. Furthermore, the semantics of all the various types of conditionals are hard to understand even for higher English proficiency learners, which means the learners make errors in comprehension and performance of the grammar. Examples of three types of the conditionals used in this study are as follows:

(1) If it rains tomorrow, we will stay home.　　　　(future conditional)

(2) If she were free, she would help you.　　　　(present-counterfactual conditional)

(3) If he had had ten million yen, he would have bought a yacht. (past-counterfactual conditional)

4.2.3 Design

During Week 1, the participants completed the pretests including three different tests, an untimed GJT, an ETT, and an EWT. In Week 2, each group, i.e., the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups, performed the ETT and received written CF. In Week 3, each group completed the same kind of ETT and received written CF again. These were the treatments the participants experienced in this study. In Week 4, the participants completed the posttests consisting of three kinds of tests, and after about 6 weeks, in Week 10 for convenience, they completed three different delayed posttests.

4.2.4 Treatment Materials and Procedure

After finishing the pretests, the first session of treatment was conducted (Appendix D). The treatment included the ETT and reception of written CF. The task consisted of six questions where the participants have to translate Japanese sentences into English in a written form. The six questions

were divided into two questions in which the participants needed to rely on the future conditional, two on the present-counterfactual conditional, and two on the past-counterfactual conditional. In case they were not able to find a base form for each verb to complete a sentence, verbs and other English vocabulary which seemed difficult for the participants to recall were put on the section named *Words* on the handout of the task in advance. The red mark of a circle and one point were given from the teacher to each correct English sentence, while only the red mark of X to each incorrect one. Errors on which the study did not focus, such as those in spelling, were not corrected.

When there was a certain error, different written CF was given to it according to the group that the learners belonged to. In the focused direct written CF group, the learners received a linguistic correct form given only to the errors relating to the linguistic category, the conditionals. For example, in the case of a sentence required for the use of the present-counterfactual conditional, whether the word *if* and simple past tense were correctly used in the *if*-clause, and whether the past tense in the auxiliary verb and a base form of a verb were precisely used in the main clause were thoroughly examined, and written CF was given only to the relevant errors. In the unfocused direct written CF group, all of the learners' errors were corrected, that is, a correct form was given to every error with the help of a native speaker of English. In the case of the focused metalinguistic written CF group, either a circle ('correct') or an X mark ('incorrect') was given to the main clause and the *if*-clause respectively in a sentence, and if 'incorrect,' the sign like *Check 1* was added around an X mark. Furthermore, in this group, the feedback sheet was distributed to the learners (Appendix E). On the sheet, the learners can find briefly summarized metalinguistic information about differences among the conditionals in addition to linguistic rules for each correct usage along with examples. They cannot, however, find the information about a correct form itself, i.e., an answer to each question. In order to avoid a risk that learners cannot find the information to self-correct, the sign like *Check 1* was placed near each error as explained in the previous chapter.

Each participant was asked to consider each error, comparing it with the information given by written CF, and subsequently (after about 10 minutes) was asked to start the next task, which means the start of the second session of treatment. The second ETT was adjusted in degree of difficulty of the first one; the number of questions, the breakdown of the questions, and sentence structures were not changed (Appendix D). Only changes in vocabulary were made. During the task, the learners were not allowed to refer to written CF again and to talk with other learners for accurate

survey of the effects of written CF. After conducting the second task, each learner's answer was scored, and the errors received written CF again that was tailored for each group.

In the next week (in Week 4), every participant joined in the session of posttests, and after about six week (in Week 10) they took three kinds of tests as a session of delayed posttests. After the delayed posttests, every participant in every group took a 50-minute English lesson, where the feedback sheets and the two handouts including answers to the ETTs were offered.

4.2.5 Testing Materials and Procedure

Three types of tests were designed for this study to measure the relative effectiveness of written CF on different kinds of tests. As a measuring tool for the effects of written CF mainly on acquisition of accurate explicit knowledge, an untimed GJT was adapted, while an ETT and an EWT were used as measuring tools for the effects of written CF on improvement in performance, that is, on the accurate use of the knowledge.

In the ETT in this study, the learners read two Japanese sentences, and then translated them into written English forms. In the EWT, they wrote an essay according to the topic they are given. Both of the tests, which seem to be typical writing activities in classroom in Japan, were adopted for the reason that they impose a different amount of cognitive load (so different effects are expected). Avoidance of the influence of TAP is taken into consideration. According to TAP theory, "we can use what we have learned if the cognitive processes that are active during learning are similar to those that are active during retrieval" (Lightbown, 2008, p. 27). In other words, the theory claims that information is best retrieved when the condition for retrieval matches the condition in which it is retrieved (Segalowitz & Lightbown, 1999). This study adopted an ETT both in the treatment and tests, and therefore was expected to have the effect of practice, to some extent, on scores in ETTs in addition to the effect of written CF. That's why another test, an EWT, was also adopted, which excluded the influence of TAP.

The untimed GJT consists of thirteen sentences that were printed on the handout. The participants were asked to judge whether there were some errors on each sentence (Appendix F). When the participants judged there were not errors, they were supposed to make a circle mark in the space indicated, while they judged there were, they made an X mark. In the case of X mark, they were then asked to underline the words or phrases that they thought included an error, and also asked

to write modified correct forms below the underlines. The procedure was thoroughly explained to the participants in advance, using the instruction on the handout to avoid a procedural error before making a linguistic error. Taking the cases when the participants who are not willing to judge and who make a circle mark on every sentence into consideration, only the ten of all thirteen sentences which clearly had errors were treated and the other three sentences which do not have errors were excluded from analysis. Focused ten sentences included four sentences for present-counterfactual conditionals, three for past-counterfactual conditionals, and three for future conditionals. One point was provided only when the learners made an X mark on each incorrect sentence and supplied a correct form, and the maximum score was ten. Around 15 minutes were assigned to this test so that every learner could fully refer to their explicit knowledge. In order to keep a balance of difficulty among the three tests, pretests, posttests, and delayed posttests, only vocabulary was changed with the sentence structures unchanged.

One of the tests developed to examine the relative effectiveness of written CF in performance, that is, on the accurate use of the knowledge, the ETT, was made up of six questions, and it was a duplicated version of the ETT in the treatment (Appendix F). Hence, six questions are divided into two on future conditionals, two on present-counterfactual conditionals, and two on past-counterfactual conditionals. In addition, English vocabulary which seemed to be difficult for the participants to recall was listed on the section named *Words* on the handout. The procedure and criteria of scoring were the same as those in the treatment. One point was given if a correct English sentence was written with adequate conditional forms, and the maximum score was six. Errors in spelling were not corrected.

The other test for analyzing the relative effectiveness of written CF on the accurate use of the knowledge was the EWT, where the learners were asked to write an essay, a short formal piece of writing dealing with a single topic, in around 60 words (Appendix F). The topics were "*If you had a special device with which you can be smaller, how would you like to use it?*" for the pretest, "*If you had a special device with which you can disappear, how would you like to use it?*" for the posttest, "*If you had a special device with which you can speak and understand any language, how would you like to use it?*" for the delayed posttest. Each topic was devised to induce the use of the present-counterfactual conditional, and was, of course, presented in Japanese to the participants to prevent English forms used in the topic from becoming a hint when writing. In scoring, the ratio of

successful use of the conditionals was calculated. If a learner used the present-counterfactual conditional twice in an essay, and one instance was correct and the other was incorrect, the ratio of correct use gained by dividing the number of successful use by the total number of conditional sentences, 0.5, was given to the learner as a score. In addition, for the sake of measuring the effects of written CF on overall accuracy in the essay writing, the number of errors per one T-unit was also calculated. T-unit is defined as "one main clause with all subordinate clauses attached to it" (Hunt, 1965, p. 20).

4.2.6 Data Analysis

The scores collected in the untimed GJT, the ETT, and the EWT through the pretest, the posttest, and the delayed posttest were subjected to a series of statistical analyses. A repeated-measures ANOVA analyzed the comparative effects of the treatment for each test score. One-way ANOVA with Holm's post hoc pair-wise comparisons were used to isolate the exact points in time where differences between the groups occurred when there was a significant Time x Group effect. Effect sizes for the ANOVA were estimated as partial eta-squared (η_p^2). Effect sizes for the pairwise comparisons were estimated using Cohen's d with values of .20, .50, and .80 indicating small, medium, and large effects, respectively (Cohen, 1988).

4.3 Results

This section first reports the relative effectiveness of three types of written CF on learners' acquisition of accurate knowledge of the conditionals measured by untimed GJTs according to their levels of English proficiency. Then, it reports the relative effectiveness of them on learners' accurate use of the knowledge measured by the ETT, and the relative effectiveness of them on overall accuracy in an essay measured by the EWT according to their English proficiency levels. Finally, it reports the comparison between the results gained in the untimed GJT and those in the EWT, focusing on improvement in the present-counterfactual conditionals. All tables of ANOVA in this study are shown in Appendix G.

Table 4.1

Descriptive Statistics for the Untimed GJT (the Conditionals, Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Focused MCF	23	5.52 (3.19)	6.35 (3.02)	6.65 (3.46)
Focused DCF	21	7.33 (2.40)	7.71 (2.29)	7.86 (1.73)
Unfocused DCF	19	5.68 (3.08)	6.42 (2.82)	7.74 (1.89)

4.3.1 Effects of written CF on the untimed GJT

4.3.1.1 Higher English Proficiency Group

Table 4.1 shows the descriptive statistics for test scores for three treatment groups (the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups) at the three different untimed GJTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there was a statistically significant effect only for Time (F (2, 120) = 8.29, p < .01, η_p^2 = .121), while there were no significant effects for Group (F (2, 60) = 1.98, *ns*, η_p^2 = .062), and for Time x Group interaction (F (4, 120) = 1.27, *ns*, η_p^2 = .041) (Figure 4.1).

Figure 4.1. Group means of the conditionals on the untimed GJT among higher proficiency learners.

Table 4.2

Descriptive Statistics for the Untimed GJT (the Present-Counterfactual Conditional, Higher Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
Focused MCF	23	2.61 (1.34)	3.04 (1.23)	2.96 (1.12)
Focused DCF	21	3.38 (1.09)	3.48 (0.91)	3.19 (0.85)
Unfocused DCF	19	2.63 (1.46)	2.37 (1.49)	2.74 (0.96)

Table 4.2 shows the descriptive statistics for test scores for three groups, focusing only on the present-counterfactual conditional at the three different testing periods to compare them with the scores in the EWT where only the present-counterfactual conditionals were treated. There were no statistically significant effects for Time x Group interaction (F (4, 120) = 1.75, *ns*, η_p^2 = .055), for Time (F (2, 120) = 0.27, *ns*, η_p^2 = .004), and for Group (F (2, 60) = 2.99, p < .10, η_p^2 = .091).

To sum up, the relative effectiveness of written CF on acquisition of accurate explicit knowledge of the conditionals (overall comprehension of three conditionals and comprehension of
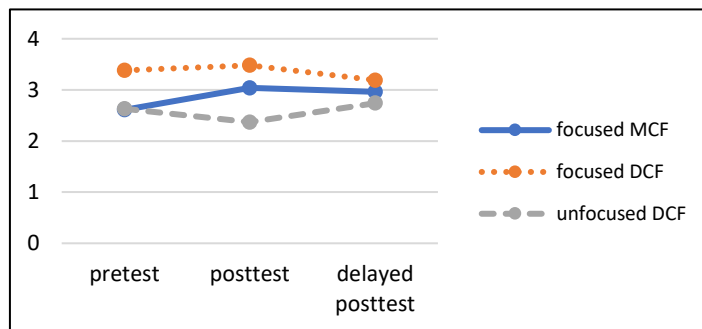
Figure 4.2. Group means of the present-counterfactual conditional on the untimed GJT among higher proficiency learners.

Table 4.3

Descriptive Statistics for the Untimed GJT (the Conditionals, Lower Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
Focused MCF	29	3.38 (3.16)	6.07 (2.48)	6.17 (2.44)
Focused DCF	27	3.37 (2.56)	4.07 (2.87)	4.04 (2.89)
Unfocused DCF	22	4.32 (2.80)	5.18 (2.67)	4.82 (2.15)

only present-counterfactual conditionals) in the higher English proficiency group were not found (Figure 4.2).

4.3.1.2 Lower English Proficiency Group

Table 4.3 shows the descriptive statistics for the test scores for three treatment groups with lower English proficiency at three timings of untimed GJTs. A repeated-measures ANOVA showed that there was no statistically significant effect for Group ($F(2, 75) = 2.13$, *ns*, $\eta_p^2 = .054$). However, there were statistically significant effects both for Time ($F(2, 150) = 21.07$, $p < .01$, $\eta_p^2 = .219$) and for Time x Group interaction ($F(4, 150) = 5.43$, $p < .01$, $\eta_p^2 = .126$). Holm pairwise comparisons
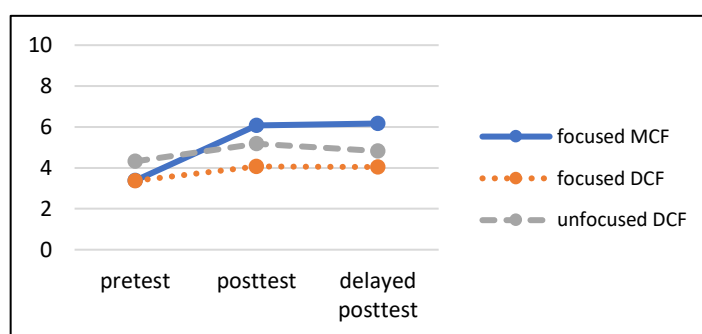
Figure 4.3. Group means of the conditionals on the untimed GJT among lower proficiency learners.

Table 4.4

Descriptive Statistics for the Untimed GJT (the Present-Counterfactual Conditional, Lower Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
Focused MCF	29	1.62 (1.40)	2.93 (0.74)	2.76 (1.19)
Focused DCF	27	1.26 (1.35)	1.81 (1.52)	1.59 (1.73)
Unfocused DCF	22	2.23 (1.44)	1.77 (1.41)	1.64 (1.15)

showed that the significant group differences were found not in the pretest but in the posttest and the delayed posttest. In the posttest, the focused metalinguistic written CF group showed a significant advantage over the focused direct written CF group with a medium effect size ($d = .75$). In the delayed posttest, the focused metalinguistic written CF group outperformed the focused direct written CF group with a large effect size ($d = .80$) (Figure 4.3).

Table 4.4 shows the descriptive statistics for test scores for three groups, focusing only on the present-counterfactual conditional at three timings of tests in order to compare them with the mean scores in the EWT where only the present-counterfactual conditional was treated. A repeated-measures ANOVA showed significant effects for Group ($F (2, 75) = 4.05$, $p < .05$, $\eta_p^2 = .096$) and for Time ($F (2, 150) = 4.61$, $p < .05$, $\eta_p^2 = .058$) and for Time x Group interaction ($F (4, 150) = 6.95$, $p < .01$, $\eta_p^2 = .156$). Holm pairwise comparisons showed that although there were no significant differences between the three groups in the pretest, there were significant differences between them in the posttest and the delayed posttest. The focused metalinguistic written CF group outperformed the focused direct written CF group ($d = .95$) with a large effect size as well as the unfocused direct written CF group ($d = 1.07$) with a large effect size in the posttest, and again, the focused metalinguistic written CF group outperformed the focused direct written CF group ($d = .79$) with a medium effect size as well as the unfocused direct written CF group ($d = .95$) with a large effect size in the delayed posttest while the difference between the focused direct written CF group and the unfocused direct written CF group in both tests did not reach statistical significance.
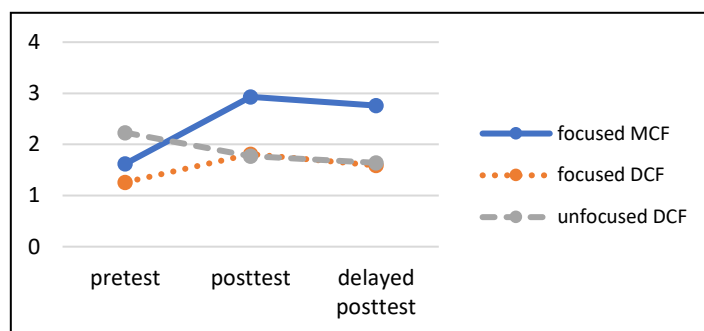
Figure 4.4. Group means of the present-counterfactual conditional on the untimed GJT among lower proficiency learners.

To recapitulate, the focused metalinguistic written CF treatment proved to be relatively effective compared with the focused and unfocused direct written CF treatments for the lower English proficiency learners for acquisition of explicit knowledge of the conditionals (overall comprehension of three conditionals and comprehension of only present-counterfactual conditionals), and focused metalinguistic written CF has a long-lasting effect (Figure 4.4).

4.3.2 Effects of Written CF on the ETT

4.3.2.1 Higher English Proficiency Group

Table 4.5

Descriptive Statistics for the ETT (the Conditionals, Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Focused MCF	23	2.17 (1.31)	4.09 (1.61)	4.04 (1.49)
Focused DCF	21	2.67 (0.84)	4.29 (1.39)	4.10 (1.31)
Unfocused DCF	19	2.74 (0.91)	3.79 (1.28)	3.63 (0.98)

Table 4.5 shows the descriptive statistics for the mean test scores for three treatment groups (the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups) at three timings of different ETTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA produced a significant effect for Time (F (2, 120) = 59.05, $p < .01$, $\eta_p^2 = .496$), with no significant effects for Group (F (2, 60) = 0.44, ns, $\eta_p^2 = .015$) and for Time x Group interaction (F (4, 120) = 1.99, $p < .10$, $\eta_p^2 = .062$). Hence, there was no significant difference between any two treatments for the effects in production measured by the ETTs (Figure 4.5).

4.3.2.2 Lower English Proficiency Group

Table 4.6 shows the descriptive statistics for the mean test scores for three treatment groups (the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups) at the three different ETTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed significant effects for Time (F (2, 150) = 51.66, $p < .01$, $\eta_p^2 = .408$), for Group (F (2, 75) = 5.11, $p < .01$, $\eta_p^2 = .120$), and for Time x Group interaction (F (4, 150) = 8.34, $p < .01$, $\eta_p^2 = .182$). Holm pairwise comparisons showed that the group differences were found not in the pretest but in the posttest and the delayed posttest. In the posttest, the focused metalinguistic written CF group ($d = .63$) and the unfocused direct written CF group ($d = .82$) significantly outperformed the focused direct written CF group with a medium effect size and a
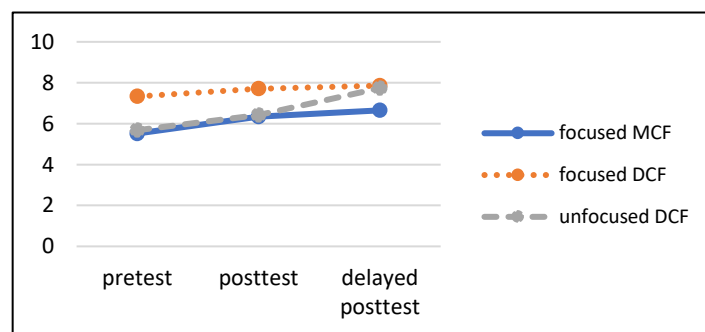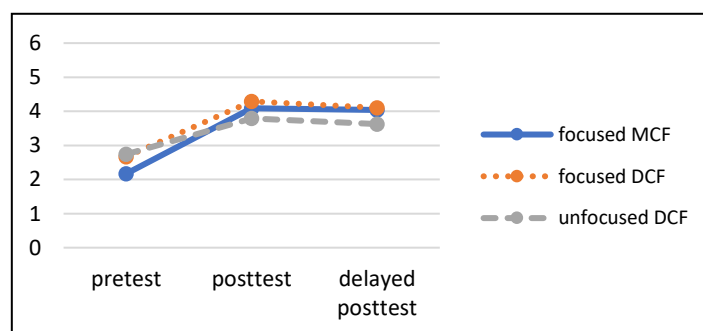

Figure 4.5. Group means of the conditionals on the ETT among higher proficiency learners.

Table 4.6

Descriptive Statistics for the ETT (the Conditionals, Lower Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
Focused MCF	29	1.86 (1.17)	3.41 (1.35)	3.17 (1.29)
Focused DCF	27	1.63 (1.06)	2.44 (1.73)	1.70 (0.90)
Unfocused DCF	22	1.73 (0.96)	3.73 (1.35)	1.82 (0.78)

large effect size respectively, but there was no significant difference between the focused metalinguistic written CF group and the unfocused direct written CF group. In the delayed posttest, the focused metalinguistic written CF group outperformed the focused direct written CF group with a large effect size ($d = 1.31$) and the unfocused direct written CF group with a large effect size ($d = 1.23$), and the significant difference between the focused direct written CF group and the unfocused direct written CF group observed in the posttest vanished. To sum up, both focused metalinguistic written CF and unfocused direct written CF proved to be effective on production for lower English proficiency learners. However, the long-lasting effect was found only in the focused metalinguistic written CF treatment (Figure 4.6).
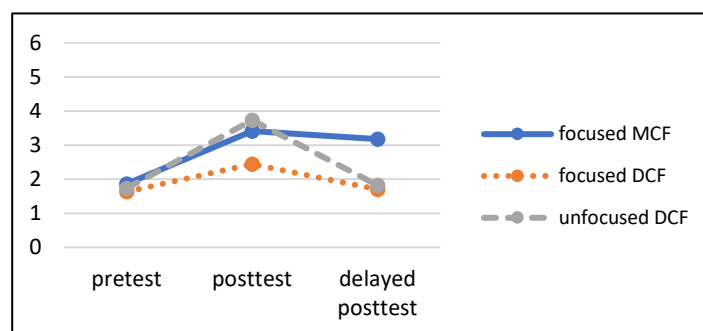
Figure 4.6. Group means of the conditionals on the ETT among lower proficiency learners.

Table 4.7

Descriptive Statistics for the EWT (the Whole Essay, Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Focused MCF	23	0.72 (0.63)	0.83 (0.62)	0.68 (0.37)
Focused DCF	21	0.47 (0.44)	0.47 (0.61)	0.54 (0.44)
Unfocused DCF	19	0.78 (0.54)	0.67 (0.56)	0.53 (0.30)

4.3.3 Effects of Written CF on the EWT

4.3.3.1 Higher English Proficiency Group

Table 4.7 shows the descriptive statistics for the number of errors per one T-unit in three treatment groups (the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups) at three timings of different EWTs (the pretest, the posttest, and the delayed posttest). Results of a repeated-measures ANOVA showed no significant effects for Time (F (2, 120) = 0.53, *ns, η_p^2* = .009), and for Group (F (2, 60) = 2.22, *ns, η_p^2* = .069), and for
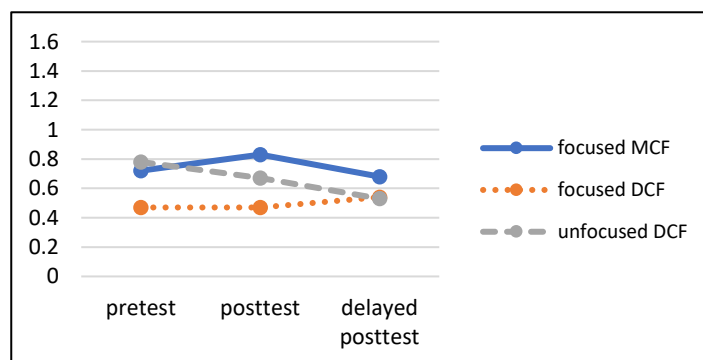
Figure 4.7. Group means of the number of errors per one T-unit on the EWT among higher proficiency learners.

Table 4.8

Descriptive Statistics for the EWT (the Present-Counterfactual Conditional, Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Focused MCF	23	0.26 (0.44)	0.76 (0.41)	0.72 (0.44)
Focused DCF	21	0.52 (0.48)	0.83 (0.36)	0.74 (0.43)
Unfocused DCF	19	0.37 (0.48)	0.55 (0.46)	0.47 (0.47)

Time x Group interaction (F (4, 120) = 0.96, *ns*, η_p^2 = .031). Therefore, there were no significant differences between any two groups for the effects on overall accuracy in an essay (Figure 4.7).

In order to compare the results of analysis in the effects of written CF on the acquisition of correct explicit knowledge of the present-counterfactual conditional measured by the untimed GJT, then, another statistical analysis was conducted focusing only on the present-counterfactual conditional in the EWT. Table 4.8 shows the descriptive statistics for the ratio of correct use of the present-counterfactual conditional in the EWTs. A repeated-measures ANOVA showed that there
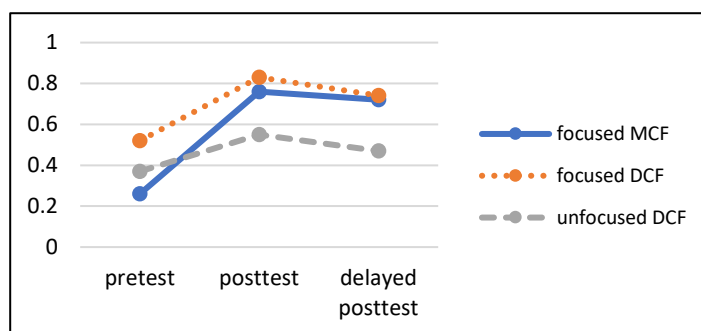
Figure 4.8. Group means of the ratio of accurate use of the conditionals on the EWT among higher proficiency learners.

Table 4.9

Descriptive Statistics for the EWT (the Whole Essay, Lower Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Focused MCF	29	1.43 (1.35)	1.12 (0.79)	0.87 (0.64)
Focused DCF	27	1.14 (0.97)	0.96 (0.81)	1.08 (0.51)
Unfocused DCF	22	0.50 (1.05)	0.80 (0.88)	0.79 (0.85)

was a statistically significant effect only for Time (F (2, 120) = 16.97, $p < .01$, $\eta_p^2 = .221$), while there were no significant effects for Group (F (2, 60) = 2.21, *ns*, $\eta_p^2 = .069$), and for Time x Group interaction (F (4, 120) = 1.81, *ns*, $\eta_p^2 = .057$), which suggested that any written CF did not show a significant advantage over other written CF on the accurate use of present-counterfactual conditionals (Figure 4.8).

4.3.3.2 Lower English Proficiency Group

Table 4.9 shows the descriptive statistics for the number of errors per one T-unit int three different treatment groups (the focused metalinguistic written CF, the focused direct written CF, and the unfocused direct written CF groups) at the three EWTs (the pretest, the posttest, and the delayed posttest). Results of a repeated-measures ANOVA showed no significant effects for Time (F (2, 150)
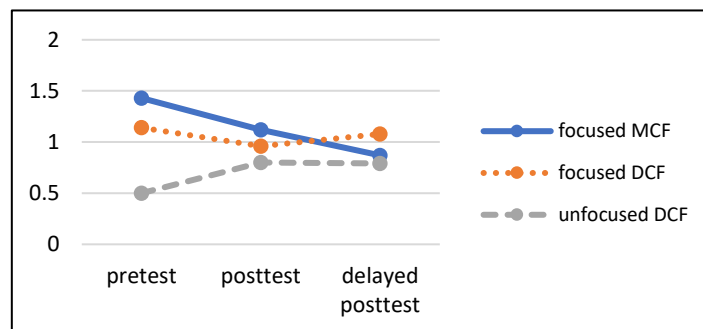
Figure 4.9. Group means of the number of errors per 1 T-unit on the EWT among lower proficiency learners.

Table 4.10

Descriptive Statistics for the EWT (the Present-Counterfactual Conditional, Lower Proficiency Group)

Groups	n	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
Focused MCF	29	0.19 (0.38)	0.50 (0.47)	0.60 (0.48)
Focused DCF	27	0.35 (0.47)	0.41 (0.46)	0.48 (0.50)
Unfocused DCF	22	0.23 (0.42)	0.58 (0.46)	0.55 (0.50)

$= 0.33$, *ns*, $\eta_p^2 = .004$), and for Time x Group interaction (F (4, 150) $= 1.92$, *ns*, $\eta_p^2 = .049$), with the main effect of Group (F (2, 75) $= 3.87$, $p < .05$, $\eta_p^2 = .094$) significant. Therefore, it is evident that there were no significant differences between any two treatment groups for the effects on overall accuracy in an essay (Figure 4.9).

In order to investigate the difference in the effects of written CF between on the acquisition of accurate explicit knowledge measured by the untimed GJT and on its accurate use, another
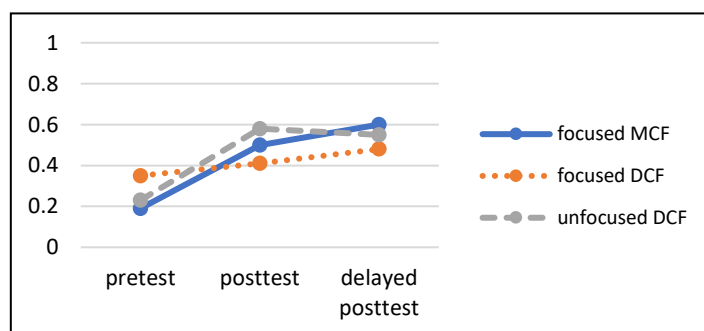
Figure 4.10. Group means of the ratio of accurate use of the conditionals on the EWT among higher proficiency learners.

Table 4.11

Summary of Data Analysis in Study 2

Category	Measuring Tool	Index of Accuracy	Higher Proficiency Group		Lower Proficiency Group	
			Time × Group	Multiple Comparisons	Time × Group	Multiple Comparisons
Comprehension	Untimed GJT	the conditionals	n.s.	—	$p<.01$	focused MCF > focused DCF [in posttest & delayed posttest]
		the preset-counterfactual conditional	n.s.	—	$p<.01$	focused MCF > focused DCF = unfocused DCF [in posttest & delayed posttest]
Performance	ETT	the conditionals	n.s.	—	$p<.01$	focused MCF = unfocused DCF > focused DCF [in posttest]; focused MCF > focused DCF = unfocused DCF [in delayed posttest]
	EWT	the whole essay	n.s.	—	n.s.	—
		the present-counterfactual conditional	n.s.	—	n.s.	—

statistical analysis was conducted, focusing only on the present-counterfactual conditional in the

EWT. Table 4.10 shows the descriptive statistics for the ratio of correct use of the present-counterfactual conditional in EWTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there was a statistically significant effect only for Time (F (2, 150) = 12.10, $p < .01$, $\eta_p^2 = .139$), while there were no significant effects for Group (F (2, 75) = 0.07, ns, $\eta_p^2 = .002$) and for Time x Group interaction (F (4, 150) = 1.46, ns, $\eta_p^2 = .038$), which suggested that any written CF did not show a significant advantage over other written CF on the correct use of present-counterfactual conditionals (Figure 4.10).

Table 4.11 summarizes the results for every comparison in this study.

4.4 Discussion

RQ 1 asked whether written CF improves accuracy in the conditionals, and RQ 2 asked whether there was any difference in the effects of written CF on the improvement in accuracy according to learners' level of English proficiency if the answer to RQ 1 was *yes*. These questions were answered by examining the results of the untimed GJT developed as a measuring tool for acquisition of accurate explicit knowledge of the conditionals, and of the ETT and the EWT developed as measuring tools for accurate use of the knowledge in performance, dividing the proficiency level into two; lower or higher. First, the relative effectiveness of written CF in the untimed GJT is discussed, which is followed by the discussion of the relative effectiveness of written CF in the ETT and the EWT.

As for the untimed GJT, the answer to RQ 1 was *yes*, but only in the lower English proficiency group. In the higher English proficiency group, it was proved that there was no significant effect for Time x Group interaction, both in the analysis of all conditionals and in the analysis of the present-counterfactual conditional, which suggests that any written CF used in this study did not result in developing the higher English proficiency learners' accurate knowledge, in other words, explicit knowledge. On the other hand, in the lower English proficiency group, the focused metalinguistic written CF group outperformed the focused direct written CF group in overall accuracy of the three conditionals in the posttest and the delayed posttest. In accuracy of the present-counterfactual conditional, the focused metalinguistic written CF group had a significant advantage over the

focused and unfocused direct written CF groups, which indicates that the provision of focused metalinguistic written CF is recommended for the lower English proficiency learners.

Considering the results of an analysis on the present-counterfactual conditional in the untimed GJT in more detail, only the unfocused direct written CF group appeared to get smaller mean scores after the treatment at the posttest, compared with the pretest. This might have been accidentally caused because the unfocused direct written CF group achieved a higher rate of accuracy in the use of the present-counterfactual conditional at the posttest than at the pretest in the EWT. However, the mean scores stayed low in the delayed posttest of the untimed GJT. This would be partly because the lower English proficiency learners in the unfocused direct written CF group, who had only an unstable knowledge of the conditionals, and who received feedback on many linguistic categories, became confused to be able to accurately judge whether the sentence is grammatical or ungrammatical, because they had to deal with three different kinds of conditionals at the same time within the allocated fifteen minutes. This suggests there was a possibility of the influence of measuring tools on the scores that resulted.

As for the ETT and the EWT, in the higher English proficiency learners, it turned out that any significant effect was observed neither in the ETT, which was almost the same as the task in the treatment, nor in the EWT, where the learners express their own opinion freely. It is assumed that the learners belonging to the higher English proficiency group, who stored a greater amount of knowledge on the conditionals than the learners in the lower English proficiency group, tended to make 'mistakes', and therefore that they were able to find the existing knowledge of the conditionals required for tests, and performed it irrespective of what kind of written CF they were provided with. As for the lower English proficiency group, the focused metalinguistic written CF group outperformed the other groups in the immediate posttest and in the delayed posttest in the ETT. Although the learners in the unfocused direct written CF group, who received feedback on a wide range of linguistic errors and raised consciousness for accuracy for them, showed significant improvement from the pretest to the posttest, the improvement disappeared in the delayed posttest, which was contrary to the focused metalinguistic written CF group. In order for the learners to get a high score on the ETT, they needed to understand the linguistic rules. That is, they needed the rules because they had no chance to use the same linguistic forms they gained by means of direct written CF in the treatment in the posttest and the delayed posttest in the ETT. The learners with a

lower level of English proficiency were considered to have little knowledge of the linguistic rules which are essential for them to deduce appropriate forms. For this reason, focused metalinguistic written CF treatment, where the learners were able to obtain metalinguistic information including rules, must have been more effective. Even if the learners in the unfocused direct written CF group immediately improved accuracy in the posttest, it faded away in the delayed posttest, after six weeks, because they did not process the feedback deeply enough to find the rules, and the rules they induced were lacking in accuracy.

No significant improvement was observed in both of the proficiency levels in another test, the EWT, which measured the effects of written CF on improvement in performance in the conditionals. Surprisingly, it was proved that unfocused direct written CF, which was provided with many linguistic errors, did not lead to significant improvement in overall accuracy in the essay. Focused metalinguistic written CF led to significant improvement in accurate knowledge and in performance measured by the ETT. However, it did not lead to any development in performance measured by the EWT. Moreover, unfocused direct written CF did not lead to improvement in accurate knowledge of the present-counterfactual conditional measured by the untimed GJT, while it led to improvement in performance related to the linguistic category measured by the ETT. Seen this way, the learners who demonstrated improvement in accuracy on knowledge level were not always able to demonstrate it at performance level, and vice versa. Furthermore, the learners who were in command of using the conditionals in some performance contexts can still make lots of errors in other contexts.

CHAPTER 5

Study 3: The Effectiveness of Written CF on the Acquisition of Explicit and Implicit Knowledge

When stating the relative effectiveness of CF strategies, we also need to compare the direct effects of written CF strategies on the acquisition of explicit knowledge, or on that of implicit knowledge. In the model of cognitive processing for L2 development through CF, what we can acquire by the stage of integration is not implicit, but explicit knowledge. There are no empirical studies confirming it in the field of written CF studies.

5.1 Research Questions

Four RQs were addressed to investigate the relative effectiveness of two types of feedback (metalinguistic written CF and direct written CF) on L2 development led by acquiring explicit and implicit knowledge of the present perfect tense according to learner's levels of proficiency (higher or lower).

RQ 1: Does written CF lead to development in implicit knowledge of the present perfect tense?
RQ 2: If so, is there any difference in the effects of written CF depending on English proficiency level?
RQ 3: Does written CF lead to development in explicit knowledge of the present perfect tense?
RQ 4: If so, is there any difference in the effects of written CF depending on English proficiency level?

5.2 Method

5.2.1 Participants

A total of 116 Japanese learners of English in high school participated in this study. They were all second-year high school students and had received at least 4 years of formal English instruction at their junior and high schools. During the first year, they freely decided their learning course of English, standard or advanced course, and they were not allowed to change their course. In this study, 52 learners (26 male and 26 female) in an advanced course are considered as being in the higher English proficiency group, and 64 learners (35 male and 29 female) in a standard course as being in the lower English proficiency group. They all took an advanced version of *GTEC for STUDENTS* by Benesse Corporation, whose maximum score is 810, before participating in this study. The means in total scores including reading, listening, and writing section were 608.1 ($SD = 69.39$) for the higher proficiency group and 498.6 ($SD = 53.08$) for the lower proficiency group. Each proficiency group was first divided into the experimental group and the control group. The experimental group was divided further into the metalinguistic written CF group and the direct written CF group. Indirect written CF, which can give negative evidence to each error and is said to generally have a smaller effect than direct written CF, was not be included in the study. The higher English proficiency group consisted of the metalinguistic written CF group ($n = 21$, $M = 607.3$, $SD = 79.04$), the direct written CF group ($n = 17$, $M = 599.3$, $SD = 62.16$), and the control group ($n = 14$, $M = 619.9$, $SD = 59.90$). On the other hand, the lower English proficiency group consisted of the metalinguistic written CF group ($n = 16$, $M = 512.0$, $SD = 55.73$), the direct written CF group ($n = 23$, $M = 508.1$, $SD = 50.30$), and the control group ($n = 25$, $M = 480.3$, $SD = 46.68$). The mean values described above are those in the total score of *GTEC for STUDENTS*.

5.2.2 Target Structure

The target structure in this study was the present perfect tense. In order to understand the present perfect tense, it is important to understand the difference in meaning between the present perfect tense and the past tense. According to Shirahata (2015), with regard to the present perfect tense, learners have difficulty in understanding the meaning that it carries rather than the form like '*have* + past participle.' The study by Aoyama (2018) showed that even most effective written CF in his study, metalinguistic written CF, was not able to improve accuracy nearer to the maximum score on the test. Based on these, the present perfect tense was chosen as a target structure in this

study.

In the study, the preset perfect sentences were divided into three types, each of which had a meaning of *completion*, *experience*, and *continuation* respectively. In the continuation type, two further different types were prepared: sentences with stative verbs and sentences with dynamic verbs. Dynamic verbs are used in progressive forms. Examples of the present perfect tense in this study are as follows:

(1) Asuka has already watched the movie.   (completion)

(2) Bob has met the singer three times before.   (experience)

(3) He has owned much money to her since 2001.   (continuation, stative verb)

     She has been painting the walls since last night.   (continuation, dynamic verb)

5.2.3 Design

During Week 1, the participants completed three kinds of pretests, i.e., an untimed grammaticality judgment test (GJT), a timed GJT, and an elicited imitation test (EIT). In Week 2, the metalinguistic written CF, the direct written CF, and the control groups performed an ETT and received different written CF strategies (of course, the control group received no written CF). In Week 3, each group completed the same kind of English translation test (ETT) and received written CF again. These were the treatments that the participants experienced in this study. In Week 4, the participants completed three kinds of posttests, and after about 6 weeks, in Week 10, they completed a series of delayed posttests including only an untimed and a timed GJT (that is, an EIT was not included). The reason why an EIT was excluded from this study will be explained later.

5.2.4 Treatment Materials and Procedure

After finishing the pretests, the first session of treatment (an ETT and provisions of written CF) was followed (Appendix H). The ETT consisted of 17 questions where the participants have to fill in the blanks with accurate verb forms, either the present perfect or past tense, using Japanese sentences or English words written outside each blank. Verbs and other English vocabulary which seemed difficult for the participants to recall were given in the section *Words* on the handout in advance. Seventeen questions were divided into six questions where the use of the past tense was

required, two questions of the present perfect tense for completion, three questions of the present perfect tense for experience, and six questions of the present perfect tense for continuation, which were further divided into three questions using stative verbs and three questions using dynamic verbs. The reason why the task contained more questions relating to continuation than those relating to completion or to experience was that the learners in the study had much difficulty in the proper use of stative and dynamic verbs. A red circle and one point were given by the teacher to each correct English sentence, while only a red X was given to each incorrect one. Errors on which the study did not focus, such as errors in spelling, were not corrected.

When some errors emerged, different kinds of written CF were given according to the group the learners belonged to. In the direct written CF group, the learners received a handout where they were able to be informed of every accurate form to each question. In the case of the metalinguistic written CF group, the feedback sheet was distributed to the learners (Appendix I). On the sheet, the learners could find brief metalinguistic information about the difference between the past tense and the present perfect tense in addition to linguistic rules for each correct use along with example sentences. In the control group, of course, no special corrective feedback was given. Each participant was asked to consider each error comparing with the information on written CF, and subsequently (after about 10 minutes), was asked to revise the first ETT. They were asked to write down corrected sentences on the handout. After confirming that every learner finished revising, the sheet for the second ETT was delivered, which meant the start of the second session of treatment. During the ETTs, the learners were not allowed to refer to written CF again or to talk with other learners. The second ETT was adjusted in degree of difficulty of the first one; the number of questions, the breakdown of the questions, and sentence structures were not changed. Only the changes in vocabulary were made. After conducting the second task, each learner's answer was scored, and they received written CF again tailored for each group.

In the next week (in Week 4) every participant joined in a series of posttests, and after about six week (in Week 10) they took part in the session of delayed posttests. After the delayed posttests, every participant in every group took a 50-minute English lesson, where the feedback sheets for the direct written CF group, the two handouts including answers to the ETTs for the metalinguistic written CF group, and both kinds of sheets for the control group were provided.

5.2.5 Testing Materials and Procedure

Three different types of tests were designed for this study to measure the effects of written CF on the acquisition of explicit knowledge and implicit knowledge. As a measuring tool for the effects of written CF on development in implicit knowledge, a timed GJT was adopted, while an untimed GJT was used as a measuring tool for the effects of written CF on that of explicit knowledge. In addition to the timed GJT, an EIT was also adopted for measuring the effects on implicit knowledge. However, in the case of the EIT, the experimental group was limited only to the metalinguistic written CF group, and the timing of tests was limited only to a pretest and a posttest.

As explained above, an EIT was also adopted in order to measure the effects of written CF on development in implicit knowledge for the reason that the construct validity of an EIT, where the effects of written CF on implicit knowledge are examined on the basis of learner's actual performance, is greater than that of a GJT, where the effects are measured on the basis of learner's comprehension (Erlam, 2006). Speakers are considered to access implicit knowledge unconsciously when they process semantic, morphological and syntactic aspects of language during tasks, such as an EIT.

The timed GJT for measuring the effects of written CF on implicit knowledge consisted of twenty-six questions (Appendix J). The participants watched and read the English sentence projected on a screen set in front of the classroom one by one, and when they judge there is no error in the sentence, they make a checkmark on '○ section' on the handout. On the other hand, when they judge there is some errors in the sentence, they make a checkmark on '× section' on the handout. The time allocated for presentation of the English sentence was calculated on the basis of the time a native speaker of English had needed to judge in a pilot study, and as a result, three or four seconds were given to each sentence (It actually took the NS one to two seconds to judge). Additional three seconds were then given for the participants to write down their answer (a checkmark) on the sheet, and a fifteen-second interval was prepared for a rest after the first thirteen questions finished. All slides on the screen were programmed in advance to change automatically according to the scheduled time. Three questions for practice were prepared for the participants to get used to this type of test. As explained earlier, learners' responses to grammatical and ungrammatical items load on separate factors, with the former tapping implicit knowledge and the latter explicit knowledge in addition to the existence of time pressure. Although the timed GJT

should keep the participants on access only to semantic processing and noticing, those who can quickly process are considered to further access to reflecting and to use explicit knowledge to identify what is incorrect and why it is incorrect. For this reason, only the sentences including no error, that is, seventeen grammatical sentences, were focused on and the rest nine ungrammatical sentences were not. In scoring, when the participant made a checkmark on the '〇 section' to a grammatical sentence, one point was given. The maximum score was seventeen.

The EIT for measuring development in implicit knowledge consists of twenty statements (Appendix J). Four out of the twenty statements were distractors and excluded from analysis (e.g. *We will get home before it will get dark*.). Analyzed sixteen statements included four statements on the past tense, four on completion, four on experience, and four on continuation of the present perfect tense. Each type of statements included two grammatical and two ungrammatical sentences. Each participant was asked to individually sit at the desk where there was a PC and to put on earphones so that he or she could concentrate on recorded English. Recorded English sentences were spoken at a normal speed by a native speaker of English. After listening, each participant orally answered a question written in Japanese shown in the PC monitor by 'yes' or 'no.' The contents of this Japanese question were related to the English statement they had read on the monitor. This was intended to maximize the possibility that they would focus on meaning rather than form of the sentence they heard and to minimize the possibility that they memorize the sentence and reproduce it with no analysis of meaning. After answering, each participant was then asked to immediately repeat the statement they heard, and when there were some errors in the statement, they had to reproduce a corrected version of it. Take the statement *Miku already passed the test* as an example. After hearing the statement, the participant orally answers the question on the screen, "Do you want to take the English test, such as *Eiken* or *TOEFL*?" by 'yes' or 'no.' Then, he or she is asked to reproduce the corrected statement like *Miku has already passed the test* or *Miku passed the test* because there is an error. Every utterance was recorded via an IC recorder, which was used for analysis later. In scoring, one point was given when the participant's response contained the correct form of the past tense or the present perfect tense. Because self-corrected utterances after the first trial may involve the use of explicit knowledge, only the first attempts were scored. Errors which were not related to the target structure were excluded from the target of scoring. As oppose to the GJTs where all participants can take the tests at one time, the EIT should be conducted individually

and needs more time. In order for the flow from the pretest to the delayed posttest to go smoothly as planned, a limited number of participants joined the EIT, and only the comparison between metalinguistic written CF and no feedback was conducted. Moreover, the timing of the test was limited only to the pretest and the posttest, not including the delayed posttest. For these reasons, the results gained from analysis of the EIT were treated complementarily to interpret the results of the timed GJT.

The untimed GJT consisted of twenty-six statements that were printed on the handout. The participants were asked to judge whether there were some errors on each sentence one by one (Appendix J). When the participants judged there were not errors, they were required to make a circle mark on the space indicated, while they judged there were, they made an X mark. In the case of an X mark, they were asked to make an underline on the words supposed to include an error, and then to write corrected forms below the underline. The procedure was thoroughly explained to the participants in advance, using the instruction on the handout. The participants were assumed to depend fully on explicit knowledge because they were given enough time to judge, but there was the possibility of using only implicit knowledge when they judge grammatical sentences as grammatical. Thus, of all twenty-six statements prepared for the timed GJT, only the seventeen statements with some grammatical error were concerned, and other nine ungrammatical sentences were excluded. Seventeen statements included eleven statements relating to the present perfect tense and six relating to the past tense. One point was given only when the learner made an X mark on each incorrect sentence and supplied a correct form, and the maximum score was seventeen. Around fifteen minutes were assigned to this test for every learner to fully refer to their explicit knowledge. In order to keep a balance of difficulty among three tests, only vocabulary was changed with the sentence structures intact. To keep the influence of TAP away, the timed and untimed GJTs, where the learners judge the grammaticality of each English sentence, and the EIT, where they listen to English and reproduce it in an oral manner were adopted.

5.2.6 Data Analysis

The scores on the timed and untimed GJTs through the pretest, the posttest, and the delayed posttest and those for the EIT through the pretest and the posttest were subjected to a series of

Table 5.1

Descriptive Statistics for the Timed GJT (Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
MCF	21	12.00 (1.48)	12.62 (1.79)	12.86 (1.78)
DCF	17	12.71 (1.90)	13.35 (1.75)	13.82 (1.62)
NF (Control)	14	12.21 (2.01)	12.86 (2.23)	13.50 (1.35)

Note. MCF = Metalinguistic written Corrective Feedback, DCF = Direct written Corrective Feedback, NF = No Feedback

statistical analyses. Repeated measures ANOVA analyzed the comparative effects of the treatment for each test score. One-way ANOVA with Holm's post hoc pair-wise comparisons were used to isolate the exact points in time where differences between the groups occurred when there was a significant Time x Group effect. Effect sizes for the ANOVA were estimated as partial eta-squared (η_p^2). Effect sizes for the pairwise comparisons were estimated using Cohen's *d* with values of .20, .50, and .80 indicating small, medium, and large effects, respectively (Cohen, 1988).

5.3 Results

This section first reports the relative effectiveness of the two types of written CF (metalinguistic written CF, direct written CF) on learners' implicit knowledge measured by the timed GJT and the EIT according to their level of English proficiency (RQs 1 and 2). Then, it reports the relative effectiveness on learners' explicit knowledge measured by the untimed GJT according to their level of English proficiency (RQs 3 and 4). All tables of ANOVA are shown in Appendix K.

5.3.1 Effects of Written CF on Implicit Knowledge

5.3.1.1 Higher English Proficiency Group in the Timed GJT
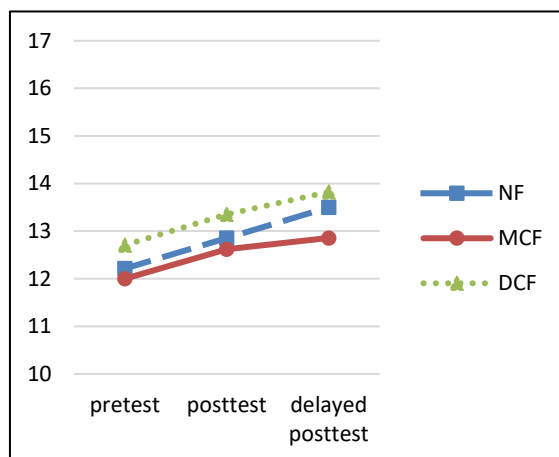
Figure 5.1. Group means on the timed GJT among higher proficiency learners.

Table 5.1 shows the descriptive statistics for test scores for three groups (the metalinguistic written CF, the direct written CF, and the control groups) at three different timed GJTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there was a statistically significant effect only for Time (F (2, 98) = 7.80, $p < .01$, $\eta_p^2 = .137$), while there were no significant effects for Group (F (2, 49) = 1.33, *ns*, $\eta_p^2 = .052$), and for Time x Group interaction (F (4, 98) = 0.13, *ns*, $\eta_p^2 = .005$) (Figure 5.1). Therefore, the effects of written CF on implicit knowledge measured by the timed GJT were not found in the higher English proficiency group.

Table 5.2

Descriptive Statistics for the Timed GJT (Lower Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
MCF	16	11.69 (2.66)	12.19 (2.01)	12.25 (2.77)
DCF	23	11.83 (1.34)	12.35 (1.52)	12.39 (1.91)
NF (Control)	25	11.20 (1.70)	11.84 (2.57)	12.24 (1.45)
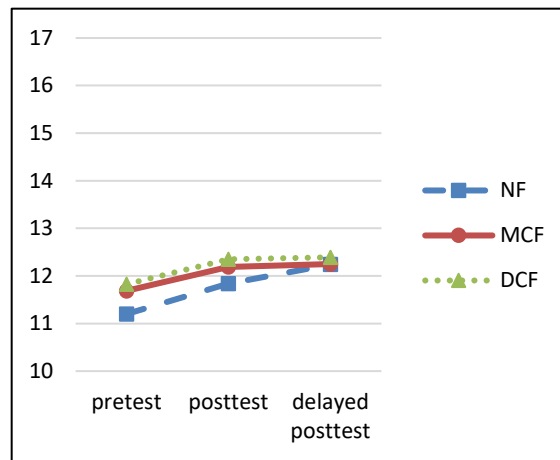
Figure 5.2. Group means on the timed GJT among lower proficiency learners.

5.3.1.2 Lower English Proficiency Group in the Timed GJT

Table 5.2 shows the descriptive statistics for test scores for three groups (the metalinguistic written CF, the direct written CF, and the control groups) at three different timed GJTs (the pretest, the posttest, and the delayed posttest) in the lower English proficiency group. Results of a repeated-measures ANOVA showed no significant effects for Time (F (2, 122) = 2.93, $p < .10$, $\eta_p^2 = .046$), for Group (F (2, 60) = 2.22, *ns*, $\eta_p^2 = .069$), and for Time x Group interaction (F (4, 120) = 0.96, *ns*, $\eta_p^2 = .031$) (Figure 5.2). Thus, the effects of written CF on implicit knowledge measured by the timed GJT were not found in the lower English proficiency group just as the higher English

Table 5.3

Descriptive Statistics for the EIT (Higher Proficiency Group)

		Pretest	Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)
MCF	15	7.93 (1.81)	8.87 (1.67)
NF (Control)	14	7.64 (2.44)	8.79 (1.61)
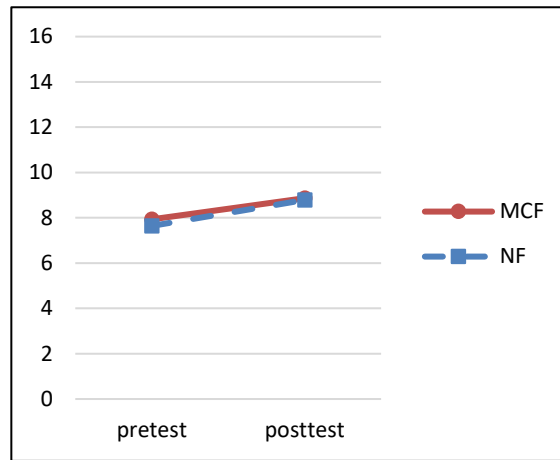
Figure 5.3. Group means on the EIT among higher proficiency learners.

proficiency group.

5.3.1.3 Higher English Proficiency Group in the EIT

Table 5.3 shows the descriptive statistics for test scores for two groups (the metalinguistic written CF group and the control group) at two different EITs (the pretest and the delayed posttest) in the higher English proficiency group. Results of a repeated-measures ANOVA revealed that there were no significant effects for Group (F (1, 27) = 0.09, *ns,* η_p^2 = .003), and for Time x Group interaction (F (1, 27) = 0.08, *ns,* η_p^2 = .003), while there was a significant effect for Time (F (1, 27)

Table 5.4

Descriptive Statistics for the EIT (Lower Proficiency Group)

Groups	*n*	Pretest Mean (*SD*)	Posttest Mean (*SD*)
MCF	14	2.36 (1.44)	3.43 (1.24)
NF (Control)	13	2.23 (1.37)	3.23 (1.05)
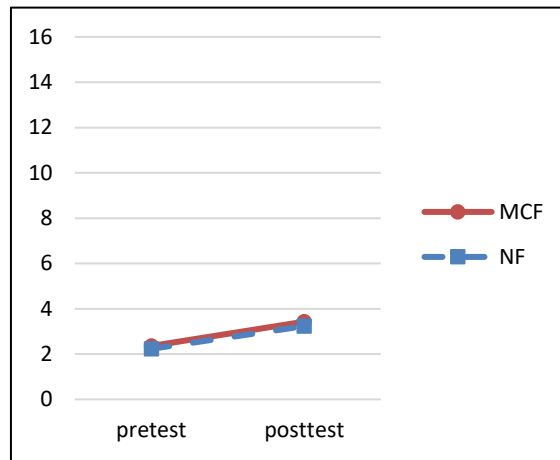
Figure 5.4. Group means on the EIT among lower proficiency learners.

$= 7.48, p < .05, \eta_p{}^2 = .217$) (Figure 5.3). Therefore, the effects of written CF on implicit knowledge measured by the EIT were not found in the higher English proficiency group.

5.3.1.4 Lower English Proficiency Group in the EIT

Table 5.4 shows the descriptive statistics for test scores for two groups (the metalinguistic written CF group and the control group) at two different EITs (the pretest and the delayed posttest) in the lower English proficiency group. Results of a repeated-measures ANOVA revealed that there were no significant effects for Group ($F (1, 25) = 0.12$, *ns*, $\eta_p{}^2 = .005$), and for Time x Group

Table 5.5

Descriptive Statistics for the Untimed GJT (Higher Proficiency Group)

		Pretest	Posttest	Delayed Posttest
Groups	*n*	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
MCF	21	6.38 (3.09)	10.67 (2.90)	10.90 (2.39)
DCF	17	7.35 (4.51)	7.94 (3.57)	7.47 (3.57)
NF (Control)	14	7.57 (3.18)	7.21 (3.59)	7.43 (2.85)

interaction (F (1, 25) = 0.03, *ns*, η_p^2 = .001), while there was a significant effect for Time (F (1, 25) = 24.00, p < .01, η_p^2 = .500) (Figure 5.4). Hence, the effects of written CF on implicit knowledge measured by the EIT were not found in lower English proficiency group.

From these results, it became clear that analyses failed to detect any significant treatment effect on development in implicit knowledge.

5.3.2 Effects of Written CF on Explicit Knowledge

5.3.2.1 Higher English Proficiency Group in the Untimed GJT

Table 5.5 shows the descriptive statistics for test scores for three groups (the metalinguistic written CF, the direct written CF, and the control groups) at three different untimed GJTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed no significant effect for Group (F (2, 49) = 1.83, *ns*, η_p^2 = .070). However, there were significant effects for Time (F (2, 98) = 18.09, p < .01, η_p^2 = .270) and for Time x Group interaction (F (4, 98) = 17.44, p < .01, η_p^2 = .416). Holm pairwise comparisons showed that the significant group differences were found not in the pretest but in the posttest and the delayed posttest. Both in the posttest and in the delayed posttest, the metalinguistic written CF group showed a significant advantage over the direct written CF group and the control group (Figure 5.5).

To sum up, the metalinguistic written CF treatment proved to be relatively effective compared
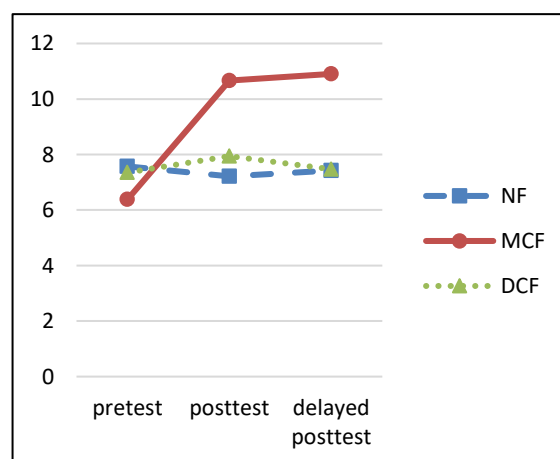
Figure 5.5. Group means on the untimed GJT among higher proficiency learners.

Table 5.6

Descriptive Statistics for the Untimed GJT (Lower Proficiency Group)

Groups	n	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
MCF	16	4.19 (3.56)	6.88 (3.30)	7.06 (3.19)
DCF	23	4.22 (3.74)	6.48 (3.01)	5.04 (2.58)
NF (Control)	25	3.76 (3.34)	4.44 (2.77)	4.72 (2.68)

with the direct written CF or no feedback treatments. Moreover, the effectiveness of the metalinguistic written CF proved to be long-lasting.

5.3.2.2 Lower English Proficiency Group in the Untimed GJT

Table 5.6 shows the descriptive statistics for test scores for three groups (the metalinguistic written CF, the direct written CF, and the control groups) at three different untimed GJTs (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed no significant effect
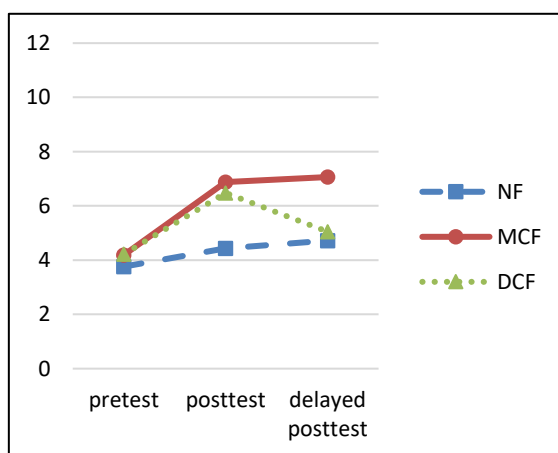
Figure 5.6. Group means on the untimed GJT among lower proficiency learners.

for Group (F (2, 61) = 1.79, ns, $\eta_p{}^2$ = .055). However, there were significant effects for Time (F (2, 122) = 26.45, p < .01, $\eta_p{}^2$ = .303) and for Time x Group interaction (F (4, 122) = 4.91, p < .01, $\eta_p{}^2$ = .139). Holm pairwise comparisons showed that the group differences were found not in the pretest but in the posttest and the delayed posttest. In the posttest, the metalinguistic written CF group (d = .82) and the direct written CF group (d = .71) significantly outperformed the control group with a large effect size and a medium effect size respectively, but there was no significant difference between the metalinguistic written CF group and the direct written CF group. In the delayed posttest, the metalinguistic written CF group outperformed the control group with a large effect size (d = .81), but there were no significant differences between the metalinguistic written CF group and the direct written CF group, and between the direct written CF group and the control group (Figure 5.6). This result showed that for lower English proficiency learners the metalinguistic written CF and the direct written CF treatments proved to be effective in a short run, but only the metalinguistic written CF treatment had a long-lasting effect.

5.4 Discussion

RQ 1 asked whether written CF affected development in implicit knowledge, and RQ 2 asked whether there was any difference in the effectiveness of written CF according to learners' English proficiency levels if the answer to RQ 1 was *yes*. These questions were answered by examining the results of the timed GJT and the EIT, dividing the proficiency level into two; lower or higher. The answer to RQ 1 was *no*. Considering the results of the timed GJTs in the pretest, the posttest, and the delayed posttest, any written CF did not result in development in implicit knowledge in both higher and lower proficiency groups, and this was also true for the result of the EIT where the timing of the test was on two levels, the pretest and the posttest. In the EIT, the metalinguistic written CF group did not outperform the control group. Hence, the study failed to illustrate the direct effectiveness of written CF on implicit knowledge, which was assumed in reactivation and reconsolidation theory from cognitive psychology.

The learner who receives direct written CF is provided not with explicit information about accurate linguistic 'rules,' but rather about 'forms.' For this reason, he or she needs to inductively

find the rules or recall them with the help of the forms, and it would not be clear whether or not the learner succeeds in really finding them, and, even if he or she succeeds, it would also not be clear whether or not the rules are correct or acceptable in the norm of L2. As a result, it would be possible for the learner to store an accurate linguistic form for each error in writing, and possibly renewed information about forms and rules at the stage of *integration*. On the other hand, in the case of metalinguistic written CF, the learner can obtain no explicit information of a correct form, but rather metalinguistic information, and therefore she or he should deduce a correct form which fits each occasion. However, it would not be clear whether the learner can really do such a thing or whether the form that the learner deduces is truly correct. It would be possible for the learner to store correct metalinguistic information and possibly renewed information about forms and rules at the integration stage. In order for the learners to be successful in the timed GJT and the EIT, they need implicit knowledge, that is, linguistic competence for them to be able to use existing knowledge stored in long-term memory automatically and instantly. They need lots of practice to transform explicit knowledge into implicit knowledge. It is for this reason that written CF, which was offered to some errors, sometimes a few errors, was not enough to develop implicit knowledge.

RQ 3 asked whether written CF affected development in explicit knowledge, and RQ 4 asked whether there was any difference in the effects of written CF according to learners' English proficiency if the answer to RQ 3 was *yes*. These questions were answered by examining the results of the untimed GJT with two different levels of proficiency; lower or higher. Different from implicit knowledge, the effects of written CF on explicit knowledge were found in both higher and lower English proficiency groups. In the higher English proficiency group, the metalinguistic written CF group had a significant advantage over the direct written CF group and the control group in both the posttest and the delayed posttest. On the other hand, in the lower English proficiency group, both the metalinguistic written CF group and the direct written CF group outperformed the control group in the posttest, while only the metalinguistic written CF group outperformed the control group in the delayed posttest.

As explained above, in order to be successful in the untimed GJT, understanding linguistic rules was required, so learners, irrespective of their English proficiency level, gained the effects of metalinguistic written CF rather than direct written CF or no feedback. With a close investigation of the errors made by the learners with a higher level of proficiency, it became clear that they tended

to make errors in the use of the progressive form for a sentence with a stative verb, and of the non-progressive form for a sentence with a dynamic verb in the present perfect structure which means continuation, although most of them made good use of the past tense and the present perfect tense in different sentences. It would be assumed that written CF, such as metalinguistic written CF, which gives a briefly summarized metalinguistic explanation of rules, is more effective than written CF, such as direct written CF, which gives linguistic forms and forces a learner to guess linguistic rules. It is difficult for a learner to deduce complicated linguistic rules only from the forms. On the other hand, the learners with a lower level of English proficiency made errors in fundamental linguistic rules as well as in complex rules, for example in making the structure *have* + the past participle. It was supposed that they were able to deduce simple and basic rules from the forms that direct written CF offered in the immediate posttest. However, mainly because the rules were not cognitively deeply analyzed or processed, they failed to become long-lasting stored rules on which they could depend in the delayed posttest.

CHAPTER 6

Study 4: The Effectiveness of Written CF according to Grammatical Item-Specific Proficiency Levels

In Studies 1 to 3, the effectiveness of written CF was investigated according to learners' levels of L2 proficiency, which means the size of learner's long-term memory store and working memory capacity that relate to both comprehension and production of L2. In order to examine the effects of proficiency in more detail, Study 4 investigated the relative effectiveness of different types of written CF, taking the learners' levels of grammatical item-specific proficiency into consideration. This grammatical item-specific proficiency was divided into three levels, i.e., higher, middle, and lower.

6.1 Research Question

The purpose of this study was to investigate the relative effectiveness of three types of written CF (direct written CF, indirect written CF and metalinguistic written CF) on an increase in accuracy in two writing tasks separately dealing with two grammatical items: present perfect and past perfect. The students were assigned into three groups according to their proficiency in each of these items.

A RQ for the study is as follows:

Is there any difference in the effects of written CF depending on grammatical item-specific proficiency levels?

6.2 Method

6.2.1 Participants

A total of 144 Japanese learners of English in high school took part in this study. They were

all first-year high school students and had received at least three years of formal English instruction at their junior high schools. After two months' experience of normal English lessons, they were supposed to decide their course of English, standard or advanced, and they were not allowed to change their course once they decided. Among the participants, there are 99 learners in an advanced course and 45 learners in a standard course. All participants took two kinds of tests which are related to target structures of the study, the present perfect tense and the past perfect tense respectively, and then, were divided into the groups according to their test scores. Ninety learners who scored from 9 to 6 points on the test of the present perfect tense were registered as a higher item-specific proficiency group. Learners who gained a maximum score, 10, were excluded from the study because there was no opportunity for provision of written CF. Fifty-six learners who scored from 9 to 3 points on the tests to the past perfect tense were registered as a middle item-specific proficiency group, while sixty learners who scored from 2 to 0 points on the same test of the past tense were listed in a lower item-specific proficiency group. That is, this study examined the effects of written CF for learners with a higher item-specific proficiency, using the grammatical item, the present perfect tense, and the effects of written CF for learners with a middle item-specific proficiency and with a lower item-specific proficiency, using the grammatical item, the past perfect tense. Both in a higher and in a middle item-specific proficiency groups, three different groups (the direct written

Grammatical item	Proficiency level	The number of participants	Score range
Present perfect	Higher	90	9－6
	Middle	3	5－3
	Lower	1	2－0
Past perfect	Higher	3	11－10
	Middle	56	9－3
	Lower	60	2－0

Figure 6.1. Participants in Study 4.

CF, the metalinguistic written CF, and the indirect written CF groups) and the control group were established, where only the direct written CF group, the metalinguistic written CF group and the control group were prepared in a lower item-specific proficiency group. This was because indirect written CF was considered almost ineffective for learners with a lower item-specific proficiency (Figure 6.1).

6.2.2 Target Structures

The target structures in this study were two grammatical items: the present perfect tense and the past perfect tense. These two items are syntactically and semantically complex and difficult for high school learners of English in Japan to understand and use accurately.

Examples of four types of the present perfect tense used in this study are as follows:

(1) I have just finished my homework.   (affirmative sentence)

(2) I have been studying English for three years.   (affirmative sentence + progressive aspect)

(3) Have you arrived in Okayama yet?   (interrogative sentence)

(4) I have never traveled by airplane.   (negative sentence)

Examples of four types of the past perfect tense used in this study are as follows:

(1) We didn't know that the lesson had been canceled.   (affirmative sentence + passive voice)

(2) I had been waiting for three hours when he appeared. (affirmative sentence + progressive aspect)

(3) Had she already gone out when you called her?   (interrogative sentence)

(4) I had not arrived in Paris until I was 40 years old.   (negative sentence)

6.2.3 Design

During Week 1, the participants completed a 120-minute English lesson where they received explicit instruction of the present perfect tense and the past perfect tense. In Week 2, they performed the pretest consisting of two different tests: the test for the present perfect tense, and the test for the past perfect tense. Then, they were divided into groups according to test scores and received written CF. Thus, completion of the pretest and the provision of written CF to errors in the pretest were

considered as the treatment in this study. In Week 4, each group completed the posttest, and after about 6 weeks, they completed the delayed posttest.

6.2.4 Treatment, Testing Materials and Procedure

The tests used in each pretest, posttest, and delayed posttest were mainly divided into two types: tests for the present perfect tense and tests for the past perfect tense. That's why two kinds of pretests were developed: the pretest for the present perfect tense and the pretest for the past perfect tense, which meant that the learners took two different types of pretests. Each of them was the English translation test (ETT) consisting of seven questions, where four questions were related to the present or past perfect tense and three questions were related to other grammatical categories that were not treated in this study (Appendix L). Around fifteen minutes were assigned to each test for every participant to fully refer to their explicit knowledge.

After finishing the two types of pretests, the participants were divided into groups according to test scores, and then given different written CF for each group. The learners who were excluded from the target of this study, for example, the learners who gained less than 5 points or more than 10 points in the pretest focusing on the present perfect tense, and the learners who gained over 10 points in the pretest focusing on the past perfect tense, received metalinguistic written CF.

Direct written CF groups received the worksheet of the ETT (the pretest) that was scored and the handout that showed every answer. Metalinguistic written CF groups received the worksheet that was scored and the feedback sheet that explained the rules of the present perfect tense or the past perfect tense and the difference between the past tense and them with some English sentences as examples (Appendix M). Indirect written CF groups received the worksheet where the errors were emphasized by underlines or marks describing insertion. The same kind of written CF was given to errors which were not focused on in the study. After receiving each written CF, even though the participants were not asked to self-correct by means of written CF (they were only asked to consider deeply why they made such errors and to try to find correct forms in their minds), many of them rewrote the sentences including some errors and added some linguistic information on their handouts. During the self-correction, they were prohibited from talking with others. The participants in the control group, of course, received no feedback.

After around 10 minutes, all participants took two types of posttest and after six weeks, they

took two types of delayed posttest. In order to keep a balance of difficulty among the tests, only vocabulary was changed without changing the sentence structures.

6.2.5 Scoring

6.2.5.1 Present Perfect Tense

Each of four questions treated in the study was scored on the basis of the criteria developed for the study. The maximum score was 10 points. Errors in spelling were not corrected. Examples of sentences and the criteria for scoring are as follows:

(1) I have just finished my homework.　　　　　　(affirmative sentence)

    2 points:　　accurate use of the present perfect tense

    1 point:　　errors on past particle

              e.g.) *I have just finish my homework.

(2) I have been studying English for three years.　　(affirmative sentence + progressive aspect)

    2 points:　　accurate use of the present perfect tense and progressive aspect

    1 point:　　errors on progressive aspect

              e.g.) *I have studied English for three years.

(3) Have you arrived in Okayama yet?　　　　　　(interrogative sentence)

    3 points:　　accurate use of the present perfect tense and interrogative expression

    2 points:　　errors on past particle or word order

              e.g.) *Have you arrive in Okayama yet?

    1 point:　　errors on past particle and word order

              e.g.) *You have arrive in Okayama yet?

(4) I have never traveled by airplane.　　　　　　(negative sentence)

    3 points:　　accurate use of the present perfect tense and negative expression

    2 points:　　errors on past particle, negative expression, or word order

              e.g.) *I have never travel by airplane.

1 point: errors on more than two categories from past particle, negative expression, and word order

e.g.) *I never have travel by airplane.

6.2.5.2 Past Perfect Tense

Each of four questions treated in the study was scored on the basis of each criteria developed for this study. The maximum score was 12 points. Errors in spelling were not corrected. Examples of sentences and the criteria for scoring are as follows:

(1) We didn't know that the lesson had been canceled. (affirmative sentence + passive voice)

3 points: accurate use of the past perfect tense and passive forms in the main clause, and of the past tense in the subordinate clause

2 points: errors on passive forms in the main clause, or on the past tense in the subordinate clause

e.g.) *We didn't know that the lesson had canceled.

1 point: errors on passive forms in the main clause, and on the past tense in the subordinate clause

e.g.) *We don't know that the lesson had cancel.

(2) I had been waiting for three hours when he appeared.

(affirmative sentence + progressive aspect)

3 points: accurate use of the past perfect tense and progressive forms in the main clause, and of the past tense in the subordinate clause

2 points: errors on progressive forms in the main clause, or on the past tense in the subordinate clause

e.g.) *I had waiting for three hours when he appeared.

1 point: errors on progressive forms in the main clause, and on the past tense in the subordinate clause

e.g.) *I had waited for three hours when he appears.

(3) Had she already gone out when you called her? (interrogative sentence)

    3 points: accurate use of the past perfect tense in the main clause and of the past tense in the subordinate clause

    2 points: errors on the past perfect tense or on word order in the main clause, or on the past tense in the subordinate clause

                  e.g.) *Had she already went out when you called for?

    1 point: errors on more than two categories from past particle, word order in the main clause, and the past tense in the subordinate clause

                  e.g.) *Had she already went out when you call her?

(4) I had not arrived in England until I was 30 years old. (negative sentence)

    3 points: accurate use of the past perfect tense, negative expression in the main clause, and the past tense in the subordinate clause

    2 points: errors on the past perfect tense, negative expression, word order in the main clause, or the past tense in the subordinate clause

    1 point: errors on more than two categories from the past perfect tense, negative expression, word order in the main clause, or the past tense in the subordinate clause

6.2.6 Data Analysis

The scores of the ETT for the present perfect tense and for the past perfect tense through the pretest, the posttest, and the delayed posttest were subjected to a series of statistical analyses. A repeated-measures ANOVA analyzed the comparative effects of the treatment for each test score. One-way ANOVA with Holm's post hoc pair-wise comparisons were used to isolate the exact points in time where differences between the groups occurred when there was a significant Time x Group effect. Effect sizes for the ANOVA were estimated as eta-squared (η^2) with values of .01, .06, and .14 indicating small, moderate, and large effects, respectively (Cohen, 1988).

6.3 Results

This section first reports the relative effectiveness of the three types of written CF (direct written CF, metalinguistic written CF, and indirect written CF) on writing tasks in higher item-specific proficiency level. Then, it reports the relative effectiveness of them in middle item-specific proficiency level. Finally, it reports the relative effectiveness of two kinds of written CF (metalinguistic written CF and direct written CF) in lower item-specific proficiency level. All tables of ANOVA are shown in Appendix N.

6.3.1 Higher Item-Specific Proficiency Level

Table 6.1 shows the descriptive statistics for test scores for three treatment groups (the direct written CF, the metalinguistic written CF, and the indirect written CF groups) and the control group (NF) at three different tests for the present perfect tense (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there was a statistically significant effect for Time (F (2, 172) = 3.21, $p < .05$, $\eta^2 = .018$), while there were no statistically significant effects for Group (F (3, 86) = 1.55, ns, $\eta^2 = .024$) and for Time x Group interaction (F (6, 172) = 1.11, ns, η^2

Table 6.1

Descriptive Statistics in Higher Item-Specific Proficiency Group

Groups	N	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
DCF	24	8.38 (0.86)	8.67 (1.28)	8.71 (1.24)
MCF	24	8.46 (0.82)	9.08 (1.11)	9.04 (1.17)
ICF	24	8.08 (1.04)	9.04 (1.57)	8.92 (2.38)
NF (Control)	18	8.39 (0.59)	8.11 (1.45)	8.28 (1.79)

Note. DCF = Direct written Corrective Feedback, MCF = Metalinguistic written Corrective Feedback, ICF = Indirect written Corrective Feedback
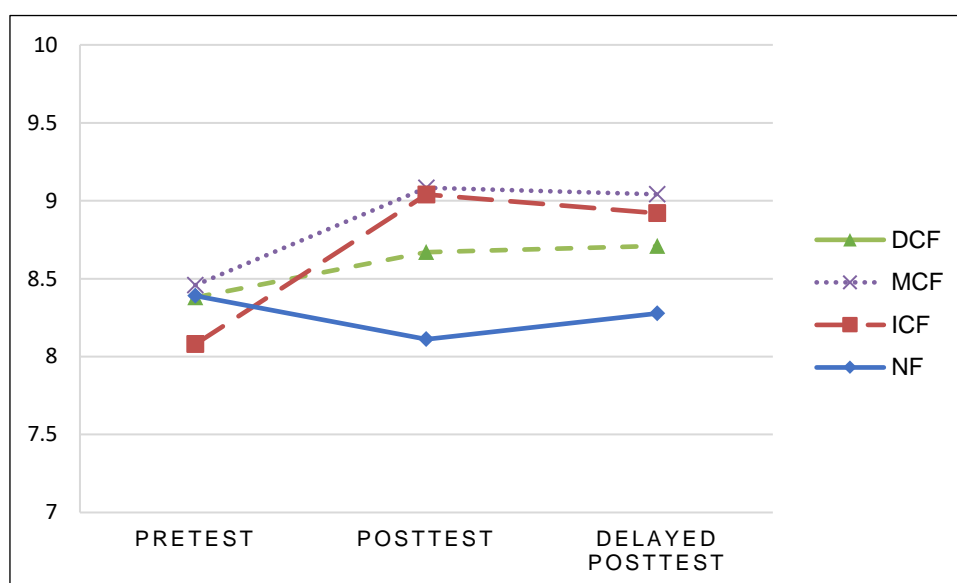
Figure 6.2. Group means in higher item-specific proficiency group.

= .019) (Figure 6.2). Therefore, the relative effectiveness of three different written CF was not identified in the learners at the higher level of item-specific proficiency.

However, since Time x Group interaction can be visually seen between the pretest and the posttest from Figure 6.2, additional repeated-measures ANOVA was conducted, with focusing only on two levels of timing, the pretest and the posttest. As a result, it showed a statistically significant effect for Time x Group interaction (F (3, 86) = 2.92, $p < .05$, $\eta^2 = .036$). Holm pairwise comparisons, however, showed that the significant group differences were found neither in the pretest nor in the posttest. Because statistically significant effects were found only in indirect written CF (F (1, 86) = 9.66, $p < .01$) and metalinguistic written CF (F (1, 86) = 4.11, $p < .01$) through the timeline (the pretest to the delayed posttest), it is possible to say that for the higher item-specific proficiency group, indirect written CF and metalinguistic written CF were more effective than direct written CF in the short run.

6.3.2 Middle Item-Specific Proficiency Level

Table 6.2 shows the descriptive statistics for test scores for three treatment groups (the direct

Table 6.2

Descriptive Statistics in Middle Item-Specific Proficiency Group

Groups	N	Pretest Mean (*SD*)	Posttest Mean (*SD*)	Delayed Posttest Mean (*SD*)
DCF	14	5.21 (1.82)	7.64 (2.38)	6.21 (2.78)
MCF	14	5.43 (2.06)	8.07 (2.87)	6.07 (3.63)
ICF	14	5.50 (1.88)	6.29 (3.08)	5.50 (3.08)
NF (Control)	14	5.50 (1.92)	5.43 (3.70)	4.79 (3.76)

written CF, the metalinguistic written CF, and the indirect written CF groups) and the control group (NF) at three different tests for the past perfect tense (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there was a statistically significant effect for Time ($F(2, 104) = 4.61$, $p < .05$, $\eta^2 = .044$), while there were no statistically significant effects for Group ($F(3, 52) = 1.24$, *ns*, $\eta^2 = .029$) and for Time x Group interaction ($F(6, 104) = 1.11$, *ns*, η^2
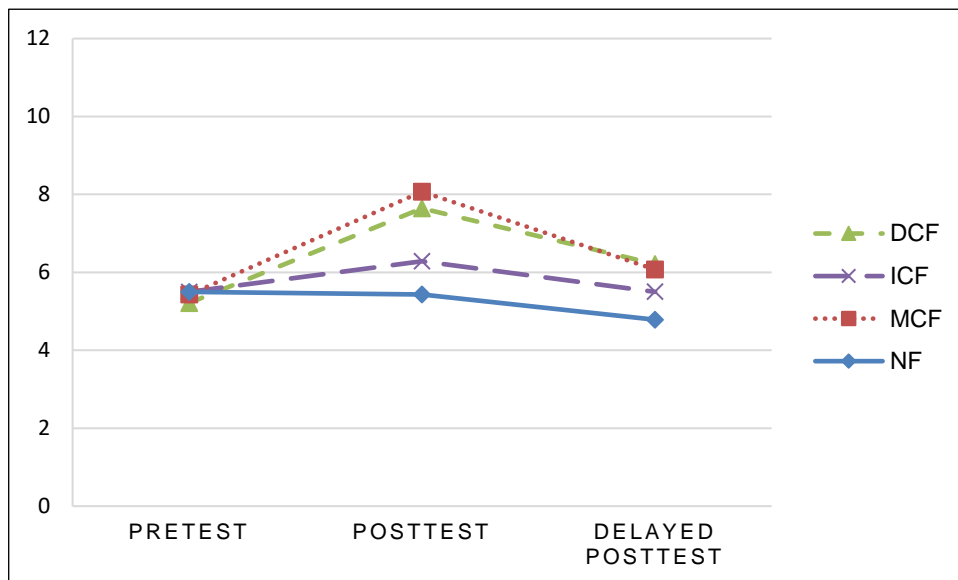
Figure 6.3. Group means in middle item-specific proficiency group.

Table 6.3

Descriptive Statistics in Lower Item-Specific Proficiency Group

		Pretest	Posttest	Delayed Posttest
Groups	N	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
DCF	20	0.60 (0.86)	3.05 (3.14)	2.95 (3.01)
MCF	20	0.60 (0.86)	6.35 (2.65)	2.75 (2.81)
NF (Control)	20	0.69 (0.73)	1.85 (1.88)	2.55 (2.62)

= .025) (Figure 6.3). Therefore, the relative effectiveness of written CF was not identified in the learners at the middle level of item-specific proficiency.

Since Time x Group interaction can be visually seen between the pretest and the posttest from Figure 6.3 just as the analysis for the higher item-specific proficiency group, additional repeated-measures ANOVA was conducted with focusing only on two levels of timing, the pretest and the posttest. As a result, it showed that there was a statistically significant effect for Time (F (1, 52) = 12.49, p < .01, η^2 = .069), while there were no statistically significant effects for Group (F (3, 52) = 0.98, *ns*, η^2 = .03) and for Time x Group interaction (F (3, 52) = 2.55, p < .01, η^2 = .042). Because statistically significant effects were found only in direct written CF (F (1, 52) = 8.80, p < .01) and metalinguistic written CF (F (1, 52) = 10.42, p < .01) through the timeline, however, it is possible to suggest that for the middle item-specific proficiency group, direct written CF and metalinguistic written CF were more effective than indirect written CF in the short run.

6.3.3 Lower Item-Specific Proficiency Level

Table 6.3 shows the descriptive statistics for test scores for two treatment groups (the direct written CF group and the metalinguistic written CF groups) and the control group at three different tests for the past perfect tense (the pretest, the posttest, and the delayed posttest). A repeated-measures ANOVA showed that there were statistically significant effects for Time (F (2, 114) = 43.18, p < .01, η^2 = .214), for Group (F (2, 57) = 4.22, p < .05, η^2 = .052) and for Time x Group interaction (F (4, 114) = 9.91, p < .01, η^2 = .098) (Figure 6.4). Holm pairwise comparisons showed
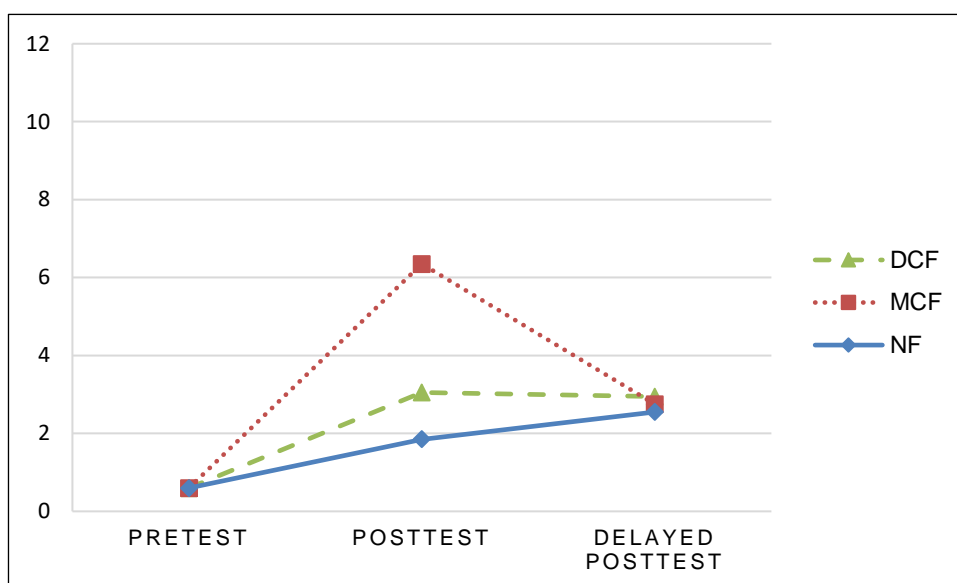
Figure 6.4. Group means in lower item-specific proficiency group.

that the significant group differences were found not in the pretest but in the posttest. However, they vanished in the delayed posttest. Only in the posttest, metalinguistic written CF showed a significant advantage over direct written CF and no feedback (MSe = 7.16, *p* < .05).

6.4 Discussion

Some researchers argue that written CF leads to development in explicit knowledge rather than implicit knowledge (Polio, 2012; Shintani et al., 2014; William, 2012). Through the treatment phase in this study, learners could acquire explicit knowledge about grammatical forms and rules. When they make some errors and receive written CF, they go into reassessment in the stage of *intake* in Figure 2.1 on Page19. If they receive indirect written CF which has only the information about the presence of an error, they have to find accurate forms or rules by themselves by means of deducing similar grammatical rules in long-term memory. With direct written CF given, learners acquire not only information about the presence of an error, but information about accurate forms.

There is no chance for them to obtain accurate rules directly from direct written CF. In the case of metalinguistic written CF, learners acquire metalinguistic information about forms and rules. However, there is no chance for them to be given accurate forms directly from metalinguistic written CF, even though there are some examples of forms on the feedback sheet. In this study the learners needed accurate linguistic rules, rather than forms to get good scores on the tests, because they could not depend on the correct forms that direct written CF provided, even if they memorized them.

In the analysis focusing on higher item-specific proficiency learners, it was proved that the differences among groups did not reach statistical significance. Although this was true for the analysis limiting the timing of the test to two levels, the pretest and the posttest, there was a tendency that metalinguistic written CF and indirect written CF were more effective than direct written CF in the short run from a detailed look at simple main effects of time. It is assumed that the learners with higher item-specific proficiency originally stored numerous explicit rules on the present perfect tense, and therefore they were able to draw the knowledge of the present perfect tense required for the tests and to use it irrespective of what kind of written CF they received. When they made 'errors', it is not clear whether they can find the rules behind forms by comparing accurate forms that direct written CF provided, with inaccurate forms they wrote. However, there was a strong possibility that they paid much attention to rules and then deduced or recalled rules they forgot by means of implicit CF such as metalinguistic written CF and indirect written CF. Therefore, they would obtain applicable rules they can use in the posttest and the delayed posttest.

As for the middle item-specific proficiency group, there were no significant differences among the groups, just like in the higher item-specific proficiency group. In the limited analysis between the pretest and the posttest, it became clear that metalinguistic written CF and direct written CF were more effective than indirect written CF in the short run. Thus, it would be safe to say that metalinguistic written CF and direct written CF are more helpful for learners with middle item-specific proficiency to gain accuracy, than indirect written CF. The learners belonging to this proficiency group, especially the learners with middle item-specific proficiency nearer to lower proficiency, would have inaccurate explicit knowledge about the past perfect tense and would make 'errors' in many cases. For them, metalinguistic written CF, which gave accurate rules and led to reconfirmation and recalling of the target item, was helpful. The learners with middle item-specific proficiency closer to higher proficiency were able to make use of information about linguistic forms

that direct written CF offered, and also to induce the underlying rules. However, indirect written CF was not enough for learners with middle item-specific proficiency to find the applicable rules that they can use in the posttest.

Finally, as for the learners belonging to the lower item-specific proficiency group, the metalinguistic written CF group significantly outperformed the direct written CF group and the control group in the posttest. However, the effectiveness of metalinguistic written CF vanished after six weeks. That is, there were no significant differences among three groups in the delayed posttest. It is assumed that the learners in this proficiency group were lacking in accurate grammatical rules of the past perfect tense and received a benefit from metalinguistic written CF that provided the rules. Even if direct written CF was given to them, they would not be able to understand the underlying rule for the posttests. However, in the delayed posttest, they would forget some grammatical rules partly because they did not process them deeply enough.

CHAPTER 7

Study 5: Learners' Attitudes toward Written CF and Text Revisions

Studies 1 to 4 in this dissertation try to identify the most effective written CF according to the learner's language proficiency level, one of the individual learner-internal cognitive factors. In Study 5, the learner's attitude toward corrective feedback strategies and text revisions is focused on, which is one of the individual learner-internal affective factors considered to influence their receptivity to error correction, and thus the effectiveness of the feedback.

7.1 Research Question

The purpose of this study is to clarify learners' attitudes toward written CF and text revisions according to learners' levels of proficiency, which is one of the learners' internal affective factors influencing the effectiveness of written CF. In addition, the relationship between the effectiveness of written CF and its preference in written CF will be examined. Studies 1 to 4 in this dissertation examined the effectiveness on the basis of learners' English proficiency, higher or lower, and thus it would be easier to understand their relationship based on the same division.

A RQ for the study is as follows:

Is there any difference in learners' attitudes toward written CF and text revisions depending on their levels of proficiency?

7.2 Method

7.2.1 Participants
The participants were the same as those in Study 2. A total of 141 Japanese learners of English

in high school participated in this study. They were all third-year high school students and had received at least 6 years of formal English instruction at their junior and high schools before attending this study. In this study, 63 learners in an advanced course are nominated as members of the higher English proficiency group, and 78 learners in a standard course as those of the lower English proficiency group. They all took an advanced version of *GTEC for STUDENTS* by Benesse Corporation, which focuses on four skills (speaking, writing, reading, and listening), and whose maximum score is 1280, before participating in this study. The means in total score were 962.7 (*SD* = 94.65) for the higher proficiency group and 814.6 (*SD* = 73.37) for the lower proficiency group. The difference in the means was statistically significant (F (1,139) = 107.98, p < .01).

7.2.2 Questionnaire

In order to examine learner attitudes toward error corrections and text revisions, the questionnaire consisting of four question items was developed. The four questions are as follows: (1) *Who do you want to correct your errors?*; (2) *How do you want your errors to be corrected?*; (3) *How many errors do you want to be corrected?*; and (4) *What do you do after receiving corrective feedback*? (See Appendix O)

Question 1: *Who do you want to correct your errors?* was set in order to examine learners' preference for a person who is in charge of providing written CF. The studies conducted by Leki (1991) and Enginarlar (1993) showed learners' stronger preference for teachers' feedback compared with peer feedback. If this preference, however, is led mainly from their trust in teachers' language proficiency, feedback from friends who have enough linguistic knowledge would be accepted by learners, especially among learners with a higher level of English proficiency. Every feedback was given by a teacher in Studies 1 to 4. Considering efficiency in classrooms, however, it would be useful to understand learners' reactions to peer feedback, which will be one of the means to reduce the effort needed for time-consuming treatment, offering written CF only from the side of a teacher.

Question 2: *How do you want your errors to be corrected?* was prepared in order to examine learners' preference for the type of written CF, i.e., indirect written CF which offers information about the presence of errors and a chance for self-correction, or direct written CF which offers both information about the presence of errors and information about accurate forms for each error with no compulsory self-correction.

Question 3: *How many errors do you want to be corrected?* was developed to examine learners' preference in the number of linguistic categories focused on in one written CF episode. That is, their preference for focused or unfocused written CF will be studied.

The last Question 4: *What do you do after receiving corrective feedback?* asked whether learners do some actions after receiving written CF. Three choices were prepared: revising the writing with written CF, only looking at their errors and written CF, or doing nothing. Written CF, which is categorized as delayed feedback, has a risk that learners ignore it, which reduces a possibility for L2 development.

7.3 Results and Discussion

7.3.1 Question 1: *Who Do You Want to Correct Your Errors?*

Table7.1 shows that the learners preferred written CF provided by a teacher to that by other students irrespective of their language proficiency, which supports the results gained in Leki (1991) and Enginarlar (1993). A chi-square test was conducted comparing the frequency of the answer in the higher proficiency group and in the lower proficiency group, but a significant difference between two groups was not found ($\chi2$ (2) = 00.00, *ns*).

The most frequent reason for their preference for written CF from the teachers' side was

Table 7.1

The Results of the First Questionnaire Item

	Options for the Answer		
Groups	Teachers	Other students	Both OK
Higher proficiency students (*N*=63)	53 (%)	0 (%)	10 (%)
Lower proficiency students (*N*=78)	60 (%)	0 (%)	18 (%)

typically "because they (teachers) are reliable" or "because their corrections are accurate". Some participants, who asked for teachers' written CF, gave reasons like "because I am not capable of correcting errors because of a lack of knowledge," or "because I was not able to correct errors accurately before," which showed their lack of confidence in correctness of error correction, showing no reference to teachers' correctness. Others chose teachers' written CF, giving various reasons like "because teachers can write some useful expressions on the sheet in addition to error corrections" or "because I would like to receive some comments and additional information about the grammar," which clearly showed that the students asked for various kinds of 'feedback' by teachers. Students who did not care about who gives written CF though that they would like to receive written CF from anyone who has correct knowledge, and also, to receive as many comments related to the contents as possible from other learners in essay writing.

The preceding studies suggest that teachers' error corrections are not consistent and are inaccurate. However, this would be solved if teachers constantly give CF on the same linguistic category for a long period. It can be assumed that students with a higher level of proficiency give accurate written CF to each other; of course, they need some training in advance. Moreover, according to the answers in the questionnaire, teachers should offer some opportunities for students to receive written CF from teachers, even if they can perform peer feedback.

7.3.2 Question 2: *How Do You Want Your Errors to Be Corrected?*

Table 7.2 shows that the learners preferred indirect written CF to direct written CF irrespective of their language proficiency, which meant that the students called for a chance that they could correct errors by themselves with some hints. Although a chi-square test was calculated comparing the frequency of the answer in the higher proficiency group and the lower proficiency group, a significant difference between two groups was not found ($\chi2$ (2) = 0.443, *ns*).

The frequent reasons for preference for direct written CF in the higher English proficiency group were typically "because it is more efficient," "indirect written CF is a waste of time," or "I want to finish writing activity with one correction episode." Frequent indirect written CF and self-correction would have become a burden for some students who did not have enough time to correct every error by themselves. On the other hand, some students in the lower English proficiency group responded like "because I cannot correct my answers even if I am given any hints for self-correction

Table 7.2

The Results of the Second Questionnaire Item

Groups	Options for the Answer		
	DCF	ICF	Both OK
Higher proficiency students (*N*=63)	21 (%)	31 (%)	11 (%)
Lower proficiency students (*N*=78)	29 (%)	34 (%)	15 (%)

through indirect written CF." Thus, they needed some hints with metalinguistic information about rules, that is, metalinguistic written CF. Students in both proficiency groups who preferred the combination of hints and self-correction stated that self-corrected words are easier to retain in minds, and direct written CF without self-correction does not lead to any development of grammar and vocabulary.

Although it became clear that the learners in the study have few opportunities for self-correction after written CF in the analysis of Question 4, giving hints for self-correction or indirect written CF, such as metalinguistic written CF, could be useful for them to have a chance to improve the situation. However, it is generally said that direct written CF is more effective for language development than indirect written CF. In order to reduce anxiety about their errors, direct written CF should be given at the final stage of a written CF episode after some provisions of indirect written CF. Future research ought to fragment indirect written CF or output-prompting CF strategies to deeply analyze which type of indirect written CF learners want.

7.3.3 Question 3: *How Many Errors (Error Categories) Do You Want to Be Corrected?*

Table 7.3 shows that almost all of the learners in both proficiency groups asked for written CF for every error, which illustrated that Japanese learners tend to worry too much over errors. Even though every error was corrected, however, few learners in the study actually rewrote their writing. A chi-square test was conducted comparing the frequency of the answer in the higher proficiency

Table 7.3

The Results of the Third Questionnaire Item

	Options for the Answer		
Groups	Unfocused	Focused	Both OK
Higher proficiency students (*N*=63)	53 (%)	3 (%)	7 (%)
Lower proficiency students (*N*=78)	64 (%)	3 (%)	11 (%)

group and the lower proficiency group, however, a significant difference between two groups was not found ($\chi 2\ (2) = 0.331$, *ns*).

Frequent answers for preference for error correction on every error were typically "because I want to aim to be perfect" or "because I cannot notice every error without written CF given to every error." Another response was typically "because every error influences the test score." Thus, integration of instruction and evaluation will be required, which can be achieved by, for example, scoring correct or incorrect use of some limited grammatical category that written CF focused on on the test. Some learners who belonged to the lower English proficiency group and who preferred focused written CF stated that they would like to focus on one grammatical category in each provision of written CF because they are not good at English and that they are sometimes confused when they receive error corrections from a wide range of grammatical categories at one time. They would like to carefully and steadily improve their skills on grammar with focused written CF.

In the preceding studies, it was proved that focused written CF, which focused on one grammatical category, was more effective for language development than unfocused written CF. This is incompatible with the result showing learners' preference for unfocused written CF, and thus learners would produce complaints if they receive only focused written CF in every writing activity. It would be effective, for example, to give unfocused written CF to errors in a short essay writing or one-sentence writing, and to give focused written CF to errors in an ETT focusing on grammar such as the counterfactual conditional or in a long essay writing. What is important must be to take

a balance between focused and unfocused written CF according to the aim of an activity.

7.3.4 Question 4: *What Do You Do after Receiving Written CF?*

A chi-square test was conducted comparing the frequency of the answer in the higher proficiency group and lower proficiency group. A significant difference between two groups was found ($\chi 2$ (2) = 7.236, *p* < .05) and effective size was middle (Cramer's *V* = .227). The detailed analysis showed that the number of learners with a lower level of proficiency who did nothing (no look and no revision) after giving written CF was significantly greater than that with a higher level of proficiency (Table 7.4).

Frequent answers in both groups for no revision were typically "because it takes lots of time to rewrite" or "because I'm busy, so I do not have enough time to rewrite." Some studies have stressed the importance of text revisions. According to Shintani et al. (2014), written CF plus the revision is more effective than written CF alone. Irrespective of whether there were multiple opportunities to revise a text with written CF (as in Chandler, 2003, and Hartshorn, Evans, Merrill, Sudweeks, Strong-Krause, & Anderson, 2010) or only a single opportunity (as in Frear's and Van Beuningen et al.'s studies), text revisions following written CF are considered to benefit greater accuracy in new writing.

Written CF plus revisions results in 'pushed output,' especially if the corrections are removed before they start to write. Swain (1985) hypothesizes that pushed output contributes to the noticing

Table 7.4

The Results of the Forth Questionnaire Item

Groups	Options for the Answer		
	Revision	Check only	Nothing
Higher proficiency students (*N*=63)	8 (%)	54 (%)	1 (%)
Lower proficiency students (*N*=78)	4 (%)	65 (%)	9 (%)

of grammatical forms that might otherwise go unattended. The need to revise involves explicit attention to the initial error and its correction, which may promote storage of the target features in long-term memory. Revisions enable learners to process written CF more deeply, helping them to consolidate their declarative or explicit knowledge of target structures. However, simply allocating additional time to process the feedback without any requirement to rewrite would not have a similar effect. Time to process the feedback and to actually rewrite would be very helpful. In addition, a number of studies have produced the results that indicate that asking learners to revise immediately after they have received feedback is advantageous. However, it became obvious that most of the learners in this study did not revise their original written texts, especially the learners in the lower English proficiency group, who were considered to need more activities for revisions, and who did not re-examine their original writing with written CF, much less revise it, compared with those in the higher English proficiency group. Therefore, teachers should prepare for some activities during a lesson for learners to revise their original texts rather than only recommend them to do so as homework. Moreover, teachers should sometimes prepare for a writing activity where learners revise their first draft again and again until no error can be found instead of starting to write on a new topic every time.

CHAPTER 8

Conclusion

8.1 Summary of the Main Findings

The main purpose of this dissertation was to identify the most effective written CF according to learners' L2 proficiency levels. In order to accomplish the purpose, the relative effectiveness of written CF was investigated by using different measuring tools. The reason why learners' proficiency levels and measuring tools should be considered is that they are assumed to influence the effectiveness of written CF. For example, it is generally said that explicit written CF strategies may be useful for relatively lower proficiency learners who need more metalinguistic explanation than higher proficiency learners. This means that a general notion that direct or explicit written CF is more effective than indirect or implicit written CF is not always true if learners' levels of English proficiency are taking into consideration. It is also said that written CF has a beneficial effect not only on text revisions but also on new pieces of writing. However, we are not sure if L2 development, which becomes manifest in text revisions, really has a positive effect on new pieces of writing because of a lack of empirical studies treating both occasions within a single research design.

From the results of Studies 1 and 2, any written CF has a positive effect on L2 development for higher proficiency learners. No written CF establishes its predominance. Study 1 investigated the relative effectiveness of direct written CF and metalinguistic written CF on text revisions and on new pieces of writing through three opportunities of providing written CF. The findings showed that there were truly positive effects of written CF on the text revision, but no clear difference was found between direct written CF and metalinguistic written CF because there was a ceiling effect, and that the relative effectiveness of written CF on new pieces of writing was not clear. In Study 2, which examined the relative effectiveness of focused direct written CF, unfocused direct written CF, and focused metalinguistic written CF on three kinds of tests; the untimed grammaticality judgment test (GJT), the English translation test (ETT), and the essay writing test (EWT), it proved that any written CF treatment led to no significant difference in improvement in accuracy. The targeted

grammatical categories in Study 2 were the same as those in Study 1, the conditionals. From these findings, any written CF is effective to higher proficiency learners in L2 development, but it is not clear whether there is a significant difference among them. This was also true for the results of Study 4, which showed no significant difference among indirect, direct, and metalinguistic written CF in long term, treating with the different grammatical categories, the present and past perfect tenses.

On the other hand, for lower proficiency learners, metalinguistic written CF, which gives learners metalinguistic information about forms and rules, seems to be the most effective for L2 development. In Study 1, metalinguistic written CF had gradual positive effects on text revisions, which then led to an increase in accuracy in the writing of new texts, while direct written CF led to improvement in accuracy only in the immediate posttest. In Study 2, in two of the three tests, the untimed GJT and the ETT, focused metalinguistic written CF proved to have a long-lasting effect. However, the predominance of metalinguistic written CF over the other types of written CF was not observed in the EWT, which was designed to exclude the influence of Transfer-Appropriate Processing (TAP). In other words, focused metalinguistic written CF was the most effective on acquisition of explicit knowledge, which was shown in the untimed GJT, and on the accurate use of the knowledge in performance, which was shown in the ETT, but the superiority of focused metalinguistic written CF over other types of written CF vanished in the EWT, another test for examining the effects on the accurate use of the knowledge in performance. In addition, focusing on the present-counterfactual conditional, the predominance of metalinguistic written CF over focused and unfocused direct written CF in the untimed GJT and in the EIT was not observed in the EWT. This result indicates that there is a gap between acquiring accurate explicit knowledge in long-term memory and its accurate use in actual performance. In the EWT, the learners needed to send their message not only accurately but also appropriately (e.g., cohesion and coherence), and needed to write multiple English sentences to construct an essay. This meant that they had to distribute their attention to many aspects of the sentence organization in an essay, rather than to correctness in each sentence in the untimed GJT and to write a few sentences in the ETT. Bitchener and Storch (2016) demonstrated that metalinguistic written CF, which offers some metalinguistic information, is more effective than any other written CF for lower English proficiency learners. However, it depends on how the effectiveness of written CF is measured, or what aspects of linguistic competence are given focus. In Study 4, which treated the grammatical categories, the

present and past perfect tenses, metalinguistic written CF proved to be more effective than direct written CF for the lower item-specific proficiency group with a large effect size in the immediate posttest. However, the effect was not long-lasting. As manifested in Study 1, metalinguistic written CF gradually improved revisions of texts positively and new written texts accordingly, and therefore a single provision of metalinguistic written CF would be insufficient in development in certain grammatical features.

The difference in effectiveness of written CF according to learners' levels of L2 proficiency can mainly arise from the relationship between the type of written CF and the quantity of explicit knowledge each learner stores in long-term memory, which then influences the quality of errors. Higher proficiency learners already store a large amount of explicit knowledge about the target grammar in long-term memory. Their errors are caused not by a complete lack of explicit knowledge, but rather by that of some small parts of the knowledge or by processing failures that arise as a result of competing plans, memory limitations, and a lack of automaticity. Irrespective of which written CF they are offered, higher proficiency learners are often able to self-correct their errors with their explicit knowledge stored in long-term memory and thus able to write errorless new texts. What they need is simply the information that signifies the presence of errors, which means every written CF is useful because it tells at least the presence of errors and is easily noticed thanks to its explicitness of written CF. In contrast, lower proficiency learners are lacking in explicit knowledge of the targeted grammar, and their errors are 'errors' in many occasions. Even if they have, their explicit knowledge is likely to be insufficient in correctness. When they receive input-providing written CF, such as direct written CF, they are likely to acquire an accurate form and renewed information about forms and rules. It is difficult in many cases for learners to induce correct rules needed for writing accurate forms in new pieces of writing, even if they can self-correct their errors in text revisions, depending on the correct rules. When they receive output-prompting written CF, such as metalinguistic written CF, they are able to reform and retest hypothesis by using metalinguistic information, and are more likely to induce correct rules which are used in text revisions and are necessary in the writing of new texts. The multi-provisions of output-prompting written CF can make the learning potential maximum by its fostering syntactic processing, another new CF provision from the interlocutor, reference to outer information resources, automatic processing, and noticing the hole. The advantage of multi-provisions of output-prompting written

CF was illustrated only in Study 1 in this dissertation, and therefore more research including a longitudinal study is needed in the future.

Study 3 tried to investigate the comparative effects of two types of written CF, i.e., direct written CF and metalinguistic written CF, on development in explicit and implicit knowledge of English present perfect tense, using three measuring tools. The findings showed that both of the tests, the timed GJT and the elicited imitation test (EIT), did not illustrate any effect of written CF on development in implicit knowledge, irrespective of which English proficiency is concerned. In contrast, the test for measuring explicit knowledge, the untimed GJT, showed the effectiveness of written CF in both higher and lower English proficiency levels. In the higher English proficiency group, only metalinguistic written CF treatment had immediate and long-lasting effects, where in the lower English proficiency group, metalinguistic written CF and direct written CF treatments had immediate effects, but only the effectiveness of metalinguistic written CF continued to stay until the delayed posttest. These results verified the validity of the information processing model claiming that the effects of written CF are displayed only in the acquisition of explicit knowledge, and are not directly exercised on development in implicit knowledge, which was expected in reactivation and reconsolidation theory in cognitive psychology. In order to develop implicit knowledge, a period of consolidation for automatization through a significant amount of practice should be necessary.

Although Kang and Han (2005) claimed that even a single provision of written CF is sufficient to improve accuracy even in the writing of new text, we should be careful about the degree of improvement. For example, it is important to clarify whether the improvement shows mastering full command of production, or merely means a slight improvement leading to decrease of some errors. It is apparent from Study 1 that metalinguistic written CF gradually improved accuracy in new pieces of writing along with improvement in decrease in errors in text revisions. Learners should be exposed to many opportunities of written CF given on the same grammatical category.

Theoretically, learners with partially developed explicit knowledge need more focused feedback because their working memory capacities are limited. Lower proficiency learners who are lacking in working memory capacities are less likely to notice CF, to reform and retest hypothesis, and to renew accurate knowledge about forms and rules in long-term memory. Too many CF strategies to various grammatical errors at one time can become heavy cognitive load for them.

However, Study 2, in which focused metalinguistic written CF, focused direct written CF and unfocused direct written CF were treated, failed to clearly prove that focused written CF is more effective than unfocused written CF as regards to an increase in accuracy.

With regard to the relative effectiveness of focused and unfocused written CF, it is very interesting that unfocused or comprehensive direct corrective feedback, which provides a correct linguistic form to every error, did not bring improvement in overall accuracy in the essay. The learners in the unfocused direct written CF group, who gained considerable feedback on misuse of the article or the third person singular present tense in the treatment, continued to make errors on the same linguistic categories in the posttest of the EWT. Because it is not clear whether the results depended on the linguistic category, and it is dangerous to overgeneralize the results gained only through the analysis of the conditionals to other linguistic categories, investigation of the effectiveness of focused written CF on a wide range of linguistic categories is needed, and the results would be helpful for language teachers, who usually give direct written CF to every error every day after lessons.

This dissertation illustrated that taking learner-internal cognitive and affective factors into consideration was important in examining the effectiveness of written CF. Except for the results that written CF did not lead to acquisition of implicit knowledge both in higher and lower proficiency learners, the relative effectiveness was highly dependent on learners' proficiency levels. This dissertation also illustrated how clarifying what kind of measuring tool was important used to examine the effectiveness of written CF. Some are text revisions where learners can directly use accurate linguistic forms given by written CF, while others are new writing tasks that ask learners to write something new where they have to find correct forms from written CF by themselves. In addition, there are different types of new writing tasks where different degree of cognitive load is placed on learners placed.

The inconsistency between effective written CF that was manifested in Studies 1 to 4 and learners' preference for written CF and text revisions was observed in some questionnaire items in Study 5. Although direct written CF is theoretically more effective than indirect written CF, the learners in Study 5 preferred indirect written CF to direct written CF, and asked for opportunities for self-correction. In addition, although focused written CF seems to be more effective in language development than unfocused written CF, the learners preferred unfocused written CF to focused

written CF, and asked for every error in every linguistic category to be corrected. This tendency of preference for indirect written CF and focused written CF was observed both in two proficiency groups. Moreover, although immediately revising texts after receiving written CF is recommended for L2 development, most of the learners did not revise their original texts, even though they compared their errors with written CF in their mind. Moreover, lower proficiency learners, compared with higher proficiency learners, were more likely to do nothing after receiving written CF.

The summary of the findings in Studies 1 to 4 is shown in Table 8.1.

8.2 Pedagogical Implications

According to Studies 1 and 2, one of the output-prompting CF strategies, metalinguistic written CF, proved to be the most effective for lower English proficiency learners, even in the writing of new texts. Moreover, it also became obvious that multiple provisions of metalinguistic written CF led to improvement in self-editing skills among learners with a lower level of proficiency. This finding means that metalinguistic written CF can lead to self-regulation, which is considered as *learning,* according to the sociocultural perspective. Learners, not always but sometimes, should be provided with opportunities for continual revisions of their original writings with metalinguistic written CF until their texts become errorless. As a result, learners can develop their editing skills, which are important to notice and self-correct their potential errors before their writings are checked by others. It is important not to give correct linguistic forms to learners' errors immediately in the first episode of error correction in order for learners to become sensitive to their own errors.

With regard to the number of grammatical categories targeted at one time, focused written CF is more effective than unfocused written CF if an increase in accuracy is concerned, and among focused strategies, focused metalinguistic written CF is more effective than focused direct written CF. The superiority of focused CF strategies over unfocused ones in lower proficiency learners would result from their limitation in their working memory capacity. Generally speaking, providing focused metalinguistic written CF over time would be useful for lower proficiency learners because they can concentrate on a limited number of grammatical categories. However, Study 5 illuminated

Table 8.1

The Summary of Studies 1 to 4

Studies	Target forms	Feedback types	Proficiency levels	Effectiveness
1	Conditionals	1. multi-DCF 2. multi-MCF 3. NF	1. Higher 2. Lower	**Higher** Revision ： — New Writing ： — **Lower** Revision ：multi-DCF & multi-MCF > NF New Writing ：multi-MCF > multi-DCF & NF

2	Conditionals	1. F_DCF
		2. U_DCF
		3. F_MCF
		(no control)
			1. Higher
			2. Lower
				Higher
				The knowledge
				  Conditionals： —
				  Present-counterfactual conditional： —
				The use
				  ETT： —
				  EWT
				   Overall accuracy： —
				   Present-counterfactual conditional： —
				Lower
				 The knowledge
				  Conditionals：F_MCF > F_DCF
				  Present-counterfactual conditional：
				    F_MCF > F_DCF & U_DCF
				The use
				  ETT： F_MCF & U_DCF > F_DCF (post)
				    F_MCF > F_DCF > U_DCF (delayed)
				  EWT
				   Overall accuracy： —
				   Present-counterfactual conditional： —

3	Present perfect	1. DCF 2. MCF 3. NF	1. Higher 2. Lower	**Higher** Implicit Knowledge： — Explicit knowledge：MCF > DCF & NF **Lower** Implicit Knowledge： — Explicit Knowledge：MCF & DCF > NF (post) MCF > DCF & NF (delayed)
4	Present & past perfect	1. ICF 2. DCF 3. MCF 4. NF	1. Higher 2. Middle 3. Lower	**Higher** ： — (ICF & MCF > DCF (pre-post)) **Middle** ： — (DCF & MCF > ICF (pre-post)) **Lower** ： MCF > DCF (post)

Note. ICF = Indirect written Corrective Feedback, DCF = Direct written Corrective Feedback, MCF = Metalinguistic written

Corrective Feedback, NF = No Feedback (Control Group), － = no significant effect for Time x Group interaction,

A > B = A has a greater effect than B

that learners ask for CF on every error in a wide range of grammatical categories. Thus, teachers

should prepare some opportunities in which they provide comprehensive feedback, unfocused written CF, even though focused written CF is considered to bring about grammatical development more than unfocused written CF from the theoretical point of view. Or teachers should provide focused written CF on a wide range of grammatical categories for a long term.

The other output-prompting CF strategy, indirect written CF, is not useful for 'errors' that are induced by a lack of correct explicit knowledge of grammatical rules. Furthermore, indirect written CF can become unessential for an increase in accuracy in the writing of new texts unless learners cannot infer a grammatical rule correctly. For this reason, indirect written CF should be given after teachers have introduced new grammatical rules through explicit instruction and after learners have acquired sufficient explicit knowledge of them. Thus, indirect written CF works well when learners make 'mistakes', not 'errors'. On the other hand, input-providing CF, direct written CF, is useful in text revisions where learners can directly use correct forms provided by the CF. However, it is unclear whether direct written CF leads to deduction of rules. In addition, direct written CF can deprive learners of the opportunity for self-correction, which is considered as important for L2 development. To conclude, it can be recommended that direct written CF and metalinguistic written CF should be provided until learners acquire sufficient explicit knowledge and until their errors become 'mistakes' and, then, indirect written CF should be offered when they make 'mistakes.' Self-correction by themselves should be also encouraged.

Because metalinguistic written CF is generally the most effective, especially for learners with a lower level of proficiency, it is also recommended that teachers should give more chance for students, who are already cognitively well developed, to be provided explicit instruction related to grammatical rules in their native tongue. As for English education today in Japan, students have a lot of opportunities to display their L2 skills on production. In order to develop fundamental skills for communication or performance, teachers should prepare activities where their students can cultivate and store accurate explicit knowledge about each grammar. Even learners with a higher level of proficiency, who generally have sufficient knowledge of grammar, cannot write English sentences without errors if they are lacking in the relevant grammatical knowledge. That is, even higher proficiency learners need the instruction which gives such information as to what is correct or what is not correct with regard to the targeted grammar. In addition, explicit instruction given to the whole class might not be enough for improving grammatical skills. Thus, after their instruction,

teachers should encourage each student to write something in English in which their lack or misunderstanding of grammar can come to the surface, and should give metalinguistic written CF to each error for each learner.

According to Study 2, overall correctness in essays did not improve in spite of providing unfocused written CF, which was given to every error that emerged on the essay. If unfocused written CF actually does not contribute to an increase in accuracy in an overall passage, teachers' time and effort to give written CF to every error would simply become a waste of time. Not only an analysis on the same conditions is called for, but new research on different conditions is required, where learners receive unfocused written CF in different types of texts.

Although it is dangerous to affirm only from Study 3, written CF is considered to give little or no influence on development in implicit knowledge. Thus, teachers should, at first, concentrate on how they can make best use of written CF for the acquisition of explicit knowledge. Implicit knowledge which learners mainly depend on when they speak is important, but it is more important to acquire explicit knowledge through written CF that becomes the foundation for development in implicit knowledge, which need a significant amount of practice. Written CF would be more effective when it is introduced together with repeated speaking activities.

Although error correction is usually conducted by teachers, recently peer correction has been introduced in classrooms. Peer correction is a method of correcting work where other students in the class correct mistakes, rather than having the teacher correct everything. However, for many L2 learners, even when they are able to notice their classmates' errors, it may be difficult to actually correct all of them due to the social and psychological nature of peer corrective feedback (Sato, 2017). Moreover, Study 5 illuminated that the learners asked for CF from a teacher, rather than from other learners, and also illuminated that they had less confidence in error correction because they believed teachers were more reliable. Sato and Lyster (2012) stated that peer corrective feedback serves dual functions to benefit both receivers' and providers' L2 development. In the process of peer corrective feedback, a learner first needs to detect an error in the input that may result in a communication breakdown or an exchange that does not involve any communication issue. In order to do so, he or she must notice the gap between the error and the accurate production. Therefore, the provider may compare the error and their interlanguage, notice that they might as well make the same error and correct it internally, and/or monitor their own CF internally and externally and,

possibly, detect the same or another error in their speech. These cognitive processes may contribute to the restructuring and consolidation of the provider's L2 knowledge. From the receiver's point of view, CF given by their peer may trigger noticing and push the learner to modify the original utterance. As Ferris (2003) argued, learners do require some training where they try to acquire the technique for peer corrective feedback, become accustomed to giving and receiving written CF with each other, and most importantly, feel confident about their skills for error correction.

To conclude, the following pedagogical instructions are recommended when written CF is provided to learners in the classroom:

(1) Lower proficiency learners should be provided with multiple provisions of metalinguistic written CF in order to improve an accuracy and an editing skill.

(2) Although focused CF strategies seem to be more effective than unfocused ones, learners demand more unfocused CF strategies than focused CF strategies. Therefore, teachers should take a balance between them.

(3) Direct written CF and metalinguistic written CF should be offered until learners acquire sufficient explicit knowledge and until their errors become 'mistakes,' and then indirect written CF should be offered when their errors are 'mistakes,' which gives a number of opportunities for self-correction.

(4) For higher proficiency learners, any CF has a positive effect on an increase in accuracy. However, this is limited to when their errors are 'mistakes.' When 'errors,' they need metalinguistic written CF just like lower proficiency learners.

(5) Written CF can be offered from both sides of a teacher and peers, but learners need sufficient training in advance in the case of peer corrective feedback.

8.3 Limitations and Recommendations for Future Research

This dissertation tried to solve the problems and overcome the shortcomings in the previous studies on written CF. Some were solved, but others were still left unexplored. In addition, new focal points that should be treated in future research emerged. As concluding remarks, limitations

and recommendations for future research will be mentioned.

8.3.1 Target Structures

This dissertation treated a limited number of grammatical categories; the conditionals, and the present and past perfect tenses, which had had little treatment in the previous studies, to the author's knowledge. However, it is still unclear whether the results gained from Studies 1 to 4 depend on targeted grammatical categories or can be generalized into others. Further written CF research should treat a wide range of grammatical categories.

Kang and Han (2015) claims that even providing written CF just once is effective. Generally speaking, learners with a higher level of proficiency, who nearly maintain constant good test scores from a posttest to a delayed posttest, are considered to already have enough explicit knowledge of the target grammar, and they can obtain not only immediate but long-lasting accuracy only with one-shot written CF. However, this was difficult for lower proficiency learners to achieve. In Study 4, every experimental group lost the rise of scores that they gained in the posttest in the delayed posttest. It is possible to suppose that this decrease of scores in the delayed posttest resulted from the complexity of the grammatical items. The past perfect tense treated in Study 1 can be thought as more difficult to understand and handle than the present perfect tense, because, for example, in the case of the past perfect tense, learners have to also understand the past expression usually accompanied with it. More research that investigates the influence of complexity of grammatical items on the effectiveness of written CF is needed.

As for the targeted grammatical categories, Study 4 had a limitation in that the same grammatical item was not targeted in every item-specific proficiency group. The study tried to investigate the effectiveness of different written CF according to three levels of grammatical item-specific proficiency, higher, middle, and lower. In order to secure the number of participants in each proficiency group, the learners with a higher item-specific proficiency were chosen on the basis of test scores for the present perfect tense, while the learners with a middle or lower item-specific proficiency were chosen on the basis of test scores for the past perfect tense. Namely, higher item-specific proficiency learners received written CF on the errors relating to the present perfect tense, and middle or lower item-specific proficiency learners received written CF on the errors relating to the past perfect tense. Because there was more or less a possibility that the difference in grammatical

items influenced the effectiveness of written CF more than proficiency levels, there is an obvious need for further research focusing on a single grammatical item.

8.3.2 Division Between Focused and Unfocused Written CF

The superiority of focused written CF over unfocused written CF, which was anticipated by Bitchener and Ferris (2012) and Ellis et al. (2008), was not clearly observed in Study 2. As mentioned earlier, the errors made by higher proficiency learners would include a significant number of 'mistakes,' rather than 'errors', which caused no significant difference between focused and unfocused written CF. Within the lower proficiency learners, those who received unfocused direct written CF on a wide range of linguistic categories, would come to naturally pay closer attention to the conditionals than to any other category, experiencing three different types of measuring tools. In addition, each question in the ETT was developed to lead the learners to use one of the three types of conditionals, and each question consisted of a few sentences, which meant that not so many error types were focused. As a result, there was a high possibility that the learners received 'less focused' written CF, rather than 'unfocused' or comprehensive written CF as in the study by Ellis et al. (2008). Future research should deal with the tasks where the participants are asked to write, for example, several sentences or an essay consisting of a few paragraphs.

8.3.3 Adoption of Various Tests and Tasks

Future research should offer various kinds of tests for measuring the effects of written CF, such as the tests for measuring the acquisition of some explicit knowledge or those for measuring the correct use of the knowledge in performance, taking the influence of TAP into consideration. It became obvious in Study 2 that learners who acquired new accurate linguistic knowledge did not always become accurate on performance. The question of what kind of additional instruction teachers should prepare for learners who have already a great amount of accurate knowledge in some grammatical rules but cannot perform it, should be examined.

The result that unfocused direct written CF did not contribute to improvement in overall accuracy in the essay deserves further investigation. The learners in the unfocused direct written CF group in Study 2, who gained a significant amount of feedback on misuse of the article or the third person singular present tense in the treatment, continued to make errors in the posttest of the EWT.

Because it is not clear that the result depends on the linguistic structure and it is dangerous to overgeneralize the result gained only through the analysis of the conditionals to other linguistic categories, the effectiveness of focused written CF on a wide range of linguistic categories needs to be examined in the future, and the results would be useful for language teachers, who mainly give direct written CF to every error on learners' written texts every day.

The studies in this dissertation used one-sentence-level translation task in each treatment. From a pedagogical perspective, essay writing tasks or paragraph writing tasks are used as often as single-sentence-level translation tasks in classroom settings. Teachers sometimes meet learners, who can use a certain grammatical rule accurately and write an errorless sentence in one-sentence writing tasks, but fail to use it accurately in essays. Future research on the effectiveness of written CF provided in various kinds of tasks should be conducted.

Although the concepts of explicit and implicit knowledge are important in SLA, it is more important for language teachers to recognize improvement in students' writing performance caused by giving them written CF. Studies would be more called for which examine the effects of written CF not only from the knowledge- or competence-based but also from the performance-based perspective.

8.3.4 Validity of Measuring Tools

Li et al. (2016), who failed to find the effect of 'oral' CF on acquisition of implicit knowledge, claimed that there was a possibility that the EIT they used did not validly measure implicit knowledge. This could be true for the EIT used in Study 3 in this dissertation. As Ellis (2005) pointed out, explicit knowledge and implicit knowledge are not two distinct competences but single continuous competence. If so, learners are accessing both kinds of knowledge in any situation, and the point is how much proportion of implicit knowledge is used or how much proportion of explicit knowledge is used. Further research should investigate the validity of timed GJTs and EITs as measuring tools for implicit knowledge.

8.3.5 Adoption of a Control Group

No control group was prepared in Study 2. This is because three different tests as measuring tools were included in the study and conducted during normal lessons, and this is because it was

preferable from educational consideration not to arrange a control group. As a result, it became unclear whether the improvement (or decline) in Study 2 was truly caused only by written CF. There is an obvious need for further research including a control group.

8.3.6 Practical Use of Written CF in Classroom

Written CF is, of course, given to learners after a writing activity. In typical English classes in Japan there must be many implicit or explicit instructions in advance so that learners can avoid making errors. Whether the effectiveness of written CF is fostered with prior instruction or not, and if so, what kind of instruction is needed should be investigated in the future. Furthermore, in addition to the effectiveness of written CF with prior instruction, the effectiveness with follow-up instruction should be paid attention to as well. Effective incorporation of written CF treatment into a natural series of lesson is important.

In usual classroom settings in Japan, learners sometimes have a question-and-answer session or a discussion in pairs or groups based on their written texts, which is one of the integrated activities focusing on more than two English skills. This means that learners have a chance to gain oral CF as well as written CF, which will make the classroom CF-rich situation. Further research should clarify what types of errors oral CF or written CF should take charge of in order for each type of CF to work and interact efficiently with each other.

8.3.7 Effects of Other Types of Feedback

Through Studies 1 to 4, the effectiveness of written CF was mainly examined on the basis of an increase in accuracy. However, feedback also plays various roles in improvement in the organization of texts and paragraphs. The quality of a written text can be determined not only by degree of correctness but also by adequate use of discourse markers, adverbs, or cohesion and coherence, for example. In order for learners to improve overall writing skills, teachers should make use of different kinds of feedback in addition to CF.

Striking the balance between content-based feedback, which is given to the contents and organization of a written text, and grammar-based feedback, which is provided to linguistic errors is important. According to the relevant studies (Kepner, 1991; Sheppard, 1992), the former is more beneficial than the latter. The only way to find a clear answer to this question is to produce more

empirical evidences treating with a wide variety of written feedback, and focusing on changes on linguistic and affective aspects written feedback brings.

8.3.8 Learner-Internal Factors

Learners' differences in attitudes or in the levels of L2 proficiency, which influence the effectiveness of written CF, should continue to be focused on. Actually, the individual learner-internal cognitive factor, L2 proficiency, affected the effectiveness of written CF, which led to conclusion that the effectiveness of written CF cannot be decided without taking learners' levels of proficiency into consideration. In conventional English lessons in Japan, each grammatical rule is taught explicitly and stored as explicit knowledge in the learner's long-term memory, which will be internalized or automatized through following enormous amount of practice. It is natural to think that each learner has a different amount of explicit knowledge and skill in utilizing it. For this reason, it would be justifiable to set up grammatical item-specific proficiency just like in Study 4, although Studies 1 to 3 adopted general L2 proficiency. The difference among higher, lower and middle levels of proficiency in this dissertation is relative, that is, there is the possibility that 'lower' proficiency in this study means 'middle' or 'higher' proficiency in other studies. Teachers should provide written CF according to learners' item-specific proficiency levels.

Many learners wait for a chance where they can correct their errors by themselves instead of being given correct answers in the first place because, they believe, self-corrected forms are more likely to stay longer in long-term memory. However, at the same time, some students complain that they have no time to do self-correction or revisions. In order to solve this problem, teachers should allot time to activities for revisions during lessons.

Learners are sensitive to their own errors and ask for unfocused or comprehensive written CF. However, learners with a lower level of proficiency tend to be confused about which error should be corrected first, and they require focused written CF strategies that are given to errors high on the list of priorities. From a viewpoint of language development, focused written CF is more advantageous than unfocused written CF. Thus, when teachers give focused written CF, it should be continuously given to a wide variety of grammatical items, which would reduce anxiety of not being corrected.

Future research should examine more direct relationships with more participants between the

tendency of preference and effectiveness of written CF, that is, how preference in written CF influences the effectiveness. Because preference could change according to participants' age, how teachers should give error correction should be determined on a basis of careful observation and investigation in their own students.

8.3.9 Reference to Sociocultural Theory

In sociocultural theory (SCT), *learning* does not mean that a learner comes to use linguistic items accurately, but means that he or she comes to depend less on assistances from other objects or other persons (Erlam, Ellis, & Batstone, 2013). It is important to note that development from SCT is evident not only in independent performance (greater accuracy on new texts) but also in a reduced reliance on assistance. The novice learner is considered independent (self-regulated) when they can write accurate texts independently, drawing on abstract representations of grammatical knowledge. SCT justifies written CF for L2 development by regarding it as one of the forms of assistance. Although the learners with a lower level of proficiency in Study 1 were not able to ultimately improve their average of test scores to around the maximum score by means of metalinguistic written CF, they were able to improve their ratio of successful self-correction. The longitudinal research should be conducted on the topic of whether learners who gradually improved the ratio of successful self-correction of some linguistic category can use it accurately without any help from others in the end.

Sheen (2011) claimed that written CF research based on SCT had not illustrated whether written CF could offer scaffolding help (finely tuned dynamic assistance in interaction) according to each learner's developmental stage. In addition, it is not clear how scaffolded knowledge becomes part of the learner's resources that the learner can deploy in independent activity. The feedback sheet developed in this dissertation, where individual learners were able to refer to the information they really needed in order to self-correct, can be one of the effective means for it.

8.3.10 Forms of Metalinguistic Written CF

Even the most effective type of written CF, especially for lower proficiency learners in this dissertation, metalinguistic written CF, has room for further improvement. The quality and quantity of metalinguistic information that appears in a feedback sheet should be tailored according to the

individual learner. The feedback sheets used in this dissertation did not include detailed information for each question, but included general information about each grammatical rule. Of course, too much information in a single feedback sheet would be inefficient. Future studies should be conducted including various types of metalinguistic written CF and examining their relative effectiveness according to grammatical item-specific proficiency in a wide range of grammatical categories.

References

Ammar, A., & Spada, N. (2006). One size fits all?: Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, *28*, 4, 543-574.

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

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

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. London: Routledge.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research, 12*, 409-431.

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

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. UK: Multilingual Matters.

Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning & Technology, 15*, 1, 41-71.

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*, 267-296.

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. New York, NY: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). San Diego, CA: Academic Press.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161-170.

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

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty and J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge:

Cambridge University Press.

Ellis, N. (2005). At the interface: How explicit knowledge affects implicit language learning. *Studies in Second Language Acquisition*, 27, 305-352.

Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning, 54,* 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). New York: Oxford University Press.

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second teaching and learning* (pp. 3-18). New York: Routledge.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*, 339-368.

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.

Enginarlar, H. (1993). Student response to teacher feedback in EFL writing. *System, 21*, 2, 193-204.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics, 27*, 464-491.

Erlam, R., Ellis, R., & Batstone, R. (2013). Oral corrective feedback on L2 writing: Two approaches compared. *System, 41,* 2, 257-268.

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*, 1-11.

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

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

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Guo, Q. (2015). The effectiveness of written CF for L2 development: A mixed-method study of

written CF types, error categories and proficiency levels. Unpublished doctoral thesis. AUT University, Auckland, New Zealand.

Gutierrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*, 423-449.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36*, 4, 543-572.

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D. I. A. N. E., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly, 44*, 1, 84-109.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (No. 3). Champaign, IL: National Council of Teachers of English.

Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–93). Harmondsworth, England: Penguin Books.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, *99*, 1-18.

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal, 75*, 3, 305-313.

Larsen-Freeman, D., & Celce-Murcia, M. (2015). *The grammar book: Form, meaning and use for English language teachers* (3rd ed.). Boston: National Geographic Learning/Cengage.

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*, 3, 203-218.

Li, S. (2010). The effectiveness of correct feedback in SLA: A meta-analysis. *Language Learning, 60*, 209-365.

Li, S., Zhu, Y., & Ellis, R. (2016). The effect of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal, 100*, 276-295.

Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. H. Han (Ed.), *Understanding second language process* (pp. 27-44). Clevedon, UK: Multilingual Matters.

Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Science 379*, 259-278.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie and T.K. Bhatia (Eds.), *Handbook of second language acquisition. Vol. 2: Second language acquisition* (pp. 413-468). New York: Academic.

Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning, 48*, 183-218.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*, 3, 399-432.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 20*, 37-66.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition, 32,* 265-302.

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

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471-497.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11*, 2, 135-128.

Nader, K. (2003). Memory traces unbound. *Trends in Neurosciences, 26*, 65-72.

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*, 203-210.

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, *21*, 375-389.

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness-practice-feedback sequence. In R.M. Dekeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*, (pp. 141-160). Cambridge: Cambridge University Press.

Sato, M. (2017). Oral peer corrective feedback: Multiple theoretical perspectives, In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second teaching and learning* (pp. 20-34). New York: Routledge.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language*

Acquisition, 34, 4, 591-626.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 2, 129-158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-31). Cambridge: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 163-182). Rowley, MA: Newbury House.

Schulz, R. A. (1996). Focus on form in the foreign language classroom: Students' and teachers' views on error correction and the role of grammar. *Foreign Language Annals*, *29*, 3, 343-364.

Schulz, R. A. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback. *The Modern Language Journal, 85*, 2, 244-258.

Segalowitz, N., & Lightbown, P. M. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics, 19*, 43-63.

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

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37*, 556-569.

Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning.* Dordrecht: Springer Netherlands.

Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 593-610). New York: Routledge.

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

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.

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*, 103-131.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Swain, M. (1985). *Communicative competence: Some roles of comprehensible input and comprehensible output in its development*. Rowley, MA: Newbury House.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*, 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*, 111-122.

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

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

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL-Review of Applied Linguistics*, *156*, 279-296.

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*, 1-41.

William, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, *21*, 321-331.

Xu, C. (2009). Overgeneralization from a narrow focus: A response to Ellis et al. (2008) and Bitchener (2008). *Journal of Second Language Writing 18*, 4, 270-275.

白畑和彦 (2015).『英語指導における効果的な誤り訂正　第二言語習得研究の見地から』大修館書店

青山聡 (2018).「日本人英語学習者の英作文の正確さに及ぼす書き言葉による訂正フィードバックの種類別効果 －文法項目に関する熟達度別に－」『教育実践学論集』第 19 号，pp.125-135.

Appendices

Appendix A: Example of the English translation test (ETT)

日本語を英語に直しなさい。　　＊「訂正→」の欄には指示があるまで何も記入しないこと。

	問　　　題	語　句		Check!
1	もしあなたの電話場号を知っていれば，電話していたのに。	電話番号	phone number	3
	If I			
訂正→	If I			
2	もっと一生懸命に取り組めば，その問題を解決できるのに。	解決する	solve	1
	If you	問題	problem	
訂正→	If you			
3	もし彼女がレポートを終えていたら，今ごろ彼女はイギリスにいるだろう。	終える	finish	4
	If she	レポート	report	
訂正→	If she	イギリス	England	
4	もし私たちがあの山の頂上にいけば，私たちはリラックスすることができる。	頂上	top	2
	If we	リラックスする	relax	
訂正→	If we			
5	もし5分早く家を出ていれば，その電車に乗ることが出来ていたのに。	5分早く	five minutes	3
	If you	乗る	catch	
訂正→	If you			
6	もし彼の電話番号を知っていれば，彼に電話できるのに。	電話する	call	1
	If I			
訂正→	If I			
7	もし明日晴れるなら，私たちはハイキングに行くことができる。	ハイキングに行く	go on a hike	2
	If it			
訂正→	If it			
8	もっと勉強に時間を費やしていたら，今ごろは夢は叶っているだろう。	費やす	spend	4
	If I	叶う	come true	
訂正→	If I			
9	もし決勝まで行っていれば，あなたが欲しいものは何でも買うのに。	決勝に行く	go to the	1
	If you			
訂正→	If you			
10	もし病気でなかったら，学校に行くことができたのに。	病気である	be sick	3
	If I			
訂正→	If I			
11	もし彼がまだテレビゲームをしているなら，彼を注意すべきだ。	テレビゲーム	the video	2
	If he	注意する	scold	
訂正→	If he			
12	もしあの時諦めていたら，あなたの会社は今ごろ成功していないだろう。	諦める	give up	4
	If you			
訂正→	If you			

Appendix B: Feedback sheet (as metalinguistic written CF)

仮定法は「**事実と異なること**」を表すものです。

■『現在』の事実と異なる場合→「仮定法過去」 ◀ **Check 1**

    If I <u>were</u> free now, I <u>would travel</u> around Japan.
      過去形　　　　　　　　助動詞の過去形＋動詞の原形

　「今ひまではないが，もしひまなら（現在の事実と異なる＝本当はひまではない！）」という場合，「仮定法過去」を用いる。If 節（従節）内は，動詞・助動詞の過去形，主節は助動詞の過去形＋動詞の限定を用いる。 "I would travel around Japan if I were free now."という語順でももちろんＯＫ。

◇直説法との違い ◀ **Check 2**
　「もし明日ひまなら，君を手伝うよ」を英語にする場合は，以下の英文ＡとＢではどちらが正しい？

  A: If I **were** free tomorrow, I **would help** you.
  B: If I **am** free tomorrow, I **will** help you.

　答えは「Ｂ」。明日ひまかどうかは不明であり，事実と異なることとは言えないので，このような場合は仮定法ではなくて，直説法を用いる。直説法であれば，<u>If節内も主節内も現在時制</u>となることに注意。

■『過去』の事実と異なる場合→「仮定法過去完了」 ◀ **Check 3**

    If I <u>had been</u> free then, I <u>would have traveled</u> around Japan.
      過去完了形　　　　　　　　助動詞の過去形＋have+過去分詞

　「あの時ひまではなかったが，もしひまだったら（過去の事実と異なる＝本当はひまではなかった）」という場合，「仮定法過去完了」を用いる。If 節（従節）内は過去完了形，主節は助動詞の過去形＋have＋過去分詞を用いる。 "I would have traveled around Japan if I had been free then."の語順でもＯＫ。

■「仮定法過去完了」＋「仮定法過去」 ◀ **Check 4**

    If I <u>had studied</u> harder, I <u>would have</u> an enjoyable university life now.
      過去完了形　　　　　　　　助動詞の過去形＋動詞の原形
  訳：「もし一生懸命勉強していたら，今ごろは楽しい大学生活を送っているだろう。」

　「もしあの時〜だったら（過去の事実と異なる），今〜だろう（現在の事実と異なる）」という場合，If 節（従節）内は過去完了形，主節は助動詞の過去形＋原形を用いる。 "I would have traveled around Japan if I had been free then."の語順でもＯＫ。

重要！　なぜ現在のことをいうのに過去形を用いるの？

仮定法で表す内容は事実と異なる話です。この現実との距離（現実⇔仮想）を英語では時制の距離（現在⇔過去）を用いて表現します。

Appendix C: ANOVA tables

ANOVA table in revision (higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	16.7989	2	8.3995	57.36	<.01	.701
Error	7.1755	49	0.1464			
Within participants						
Time	0.0435	2	0.0218	0.42	*ns*	.008
Time x Group	0.4770	4	0.1192	2.28	<.10	.085
Error	5.1259	98	0.0523			
Total	29.6208	155				

ANOVA table in revision (lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	4.0803	2	12.0402	156.39	<.01	.850
Error	4.2344	55	0.0770			
Within participants						
Time	0.0718	2	0.0359	1.47	*ns*	.026
Time x Group	0.7025	4	0.1756	7.20	<.01	.207
Error	2.6840	110	0.0244			
Total	31.7730	173				

ANOVA table in new writing (higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	586.5074	2	293.2537	5.48	<.01	.182
Error	2619.8458	49	53.4662			
Within participants						
Time	4.1030	2	2.0515	0.38	*ns*	.008
Time x Group	81.9747	4	20.4937	3.80	<.01	.134
Error	528.1095	98	5.3889			
Total	3820.5404	155				

ANOVA table in new writing (lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	536.7746	2	268.3873	4.22	<.05	.133
Error	3500.0001	55	63.6364			
Within participants						
Time	136.4339	2	68.2169	9.62	<.01	.149
Time x Group	97.5080	4	24.3770	3.44	<.05	.111
Error	780.0975	110	7.0918			
Total	5050.8141	173				

Appendix D: Examples of the English translation task (ETT)

Test 1

以下の日本語を表す英語を書きなさい。

1	「もしあと一年の命だといわれたら，どうする？」「そうだなあ，まず自分にとって何が一番大切なのか，よく考えるだろうなあ」

Words: one year to live 「１年の命」

2	教員の過重労働が話題になっている。より多くの日本人たちが，この社会問題にもっと早くから注意を向けてくれていたらよかったのに。

Words: overworking 「過重労働」、topic「話題」、social problem「社会問題」

3	日本も徐々にカード社会に変わりつつある。このごろはプリペイドカードを持っていると，乗り物に乗り，食事をして，買い物をすることすらできてしまう。

Words: a cashless society「カード社会」、a prepaid card 「プリペイドカード」、take buses or trains「乗り物に乗る」

Test 2

以下の日本語を表す英語を書きなさい。

1	あなたがある人をあるがままに扱えば，そのままで変わらないが，あるべき理想の姿として扱えば，おそらく彼はそのような人になってゆくだろう。

Words: as S is/are 「Sのあるがまま」, remain the same「そのままで変わらない」、as ideals 「理想の姿として」

2	子どもの時代にあなたはおそらく空高く鳥の飛ぶのをたびたび眺めたでしょう。「もし鳥なら空を飛べるのに」と思ったに違いありません。

Words: high up in the sky「空高く」

3	５００年前のヨーロッパには男のロマンがあふれている。私は船乗りになって、世界の果てを目指して冒険に出たことでしょう。

Words: Europe 500 years ago「500年前のヨーロッパ」、man's dream 「男のロマン」、
be full of「～であふれている」、sailar「船乗り」、the end of the world「世界の果て」

Appendix E: Feedback sheet (as metalinguistic written CF)

仮定法は「**事実と異なること**」を表すものです。

- -

■『現在』の事実と異なる場合→「仮定法過去」　　`Check 1`

　　If I <u>were</u> free now, I <u>would travel</u> around Japan.
　　　　過去形　　　　　　　　助動詞の過去形＋動詞の原形

　「今ひまではないが，もしひまなら（現在の事実と異なる＝本当はひまではない！）」という場合，「仮定法過去」を用いる。If節（従節）内は，動詞・助動詞の過去形，主節は助動詞の過去形＋動詞の原形を用いる。 "I would travel around Japan if I were free now." という語順でももちろんＯＫ。

＊現在の事実と異なる願望を表すときは，"I wish ＋ 仮定法過去"を用いる。　`Check 2`

　　I wish I <u>could</u> live in London. 「ロンドンに住めたらいいのになあ」
　　　　　　　　過去形

◇直説法との違い　`Check 3`

　「もし明日ひまなら，君を手伝うよ」を英語にする場合は，以下の英文ＡとＢではどちらが正しい？

　A:　　If I **were** free tomorrow, I **would help** you.
　B:　　If I **am** free tomorrow, I **will** help you.

　答えは「Ｂ」。明日ひまかどうかは不明であり，「事実と異なる」とは言えないので，このような場合は仮定法ではなくて，直説法を用いる。直説法であれば，<u>If節内は現在時制</u>になることに注意。

■『過去』の事実と異なる場合→「仮定法過去完了」　`Check 4`

　　If I <u>had been</u> free then, I <u>would have traveled</u> around Japan.
　　　　過去完了形　　　　　　　　助動詞の過去形＋have＋過去分詞

　「あの時ひまではなかったが，もしひまだったら（過去の事実と異なる＝本当はひまではなかった）」という場合，「仮定法過去完了」を用いる。If節（従節）内は過去完了形，主節は助動詞の過去形＋have＋過去分詞を用いる。 "I would have traveled around Japan if I had been free then." の語順でもＯＫ。

＊過去の事実と異なる願望を表すときは，"I wish ＋ 仮定法過去完了"を用いる。　`Check 5`

　　I wish I <u>had been</u> more careful. 「もっと注意していたらよかったのになあ」
　　　　　　　過去完了

重　要　　**現実との距離＝時制の距離！**

仮定法で表す内容は事実と異なる話です。この現実との距離（現実⇔仮想）を英語では時制の距離（現在の事実と異なる場合は過去，過去の事実と異なる場合は過去完了）を用いて表現します。

Appendix F: Examples of the tests

The untimed grammaticality judgement test (untimed GJT)

問題No.	○/×	
例	○	私は昨日プールで泳ぎを楽しみました。 I enjoyed swimming in the pool yesterday.
例	×	私はこないだの夏にアメリカに行った。 I go to America last summer. 　went
1		もう少し早く出ていれば，ラッシュアワーを避けることが出来ていたのに。 If I had left earlier, we could avoid the rush hour.
2		彼女がまだ寝ていたら，私はあなたに彼女を起こしてもらいたい。 If she was still sleeping, I want you to wake her up.
3		私が間違っていれば、謝るのに。 If I were wrong, I would apologize.
4		勉強すれば，彼はいい成績が取れるのに。 If he had studied, he would get good marks.
5		彼が本当の科学者なら，そのようには考えないだろう。 If he had been a true scientist, he would not think that way.
6		事前に知っていれば，その会議に参加していただろうに。 If I had known in advance, I would attend the meeting.
7		インドに住んでいれば，そこの全ての食べ物を食べるのに。 If I had lived in India, I would try all the foods there.
8		もしその時風邪を引いていなかったら，あなたと釣りに行っていたのに。 If I had not had a cold then, I would have gone fishing with you.
9		もしもっとお金があれば，そのバッグを買っていたのに。 If I had had more money, I could buy the bag.
10		この鍵を使えば，そのドアを開けることができるよ。 If you use this key, you can open the door.
11		彼女がコンピュータを使えるなら，容易に仕事を見つけられるのに。 If she had been able to use a computer, she could easily find a job.
12		トムが明日来るなら，私は彼を公園に連れていく。 If Tom will come tomorrow, I will take him to the park.
13		天気が良ければ，パーティは成功していたのに。 If the weather had been good, the party would be successful.

141

The English translation test (ETT)

以下の日本語を表す英語を書きなさい。

| 1 | 問題分析や議論の方法を教わっていないのであれば，「問題分析ができない」とか「自分の意思がない」と学生たちを批判することはできない。 |

Words: how to analyze or discuss problems「問題分析や議論の方法」、
criticize「批判する」、opinion「意思」

| 2 | もっと若いときのことだったら両親に話しもしただろう。しかし 四十近くになってからの同棲，結婚なので，事後報告ということで済ませた。 |

Words: marriage「結婚」、〜と結婚する「marry〜」「get married」、
almost 40 years old「四十近く」、after the fact「事後報告」

| 3 | 学校の先生が教科書に書いてあることしか知らなければ，生徒の質問に答えることはできないだろう。だから先生たちには，目の前にいる生徒同様に，学び続ける必要がある。 |

Words: textbook「教科書」

The essay writing test (EWT)

以下のテーマで，６０語程度の英文を書きなさい。

● もしドラえもんのひみつ道具「*ほんやくコンニャク (honyac konjac)」があれば，どのように利用
したいですか。
＊これを食べると，あらゆる言語が母国語に翻訳されて聞こえる。自分が話す言葉は相手の母国語に
なる。文章も翻訳して読み取れる。

■3年（　　　）組（　　　）番　発　・　標　　名前（　　　　　　　　　　　　　）

（＿＿＿＿＿）words

Appendix G: ANOVA tables

ANOVA table in the untimed GJT (the conditionals, higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	70.3503	2	35.1752	1.98	*ns*	.062
Error	1064.0250	60	17.7337			
Within participants						
Time	47.8395	2	23.9197	8.29	<.01	.121
Time x Group	14.6418	4	3.6605	1.27	*ns*	.041
Error	346.0938	120	2.8841			
Total	1542.9504	188				

ANOVA table in the untimed GJT (the present-counterfactual conditional, higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	18.9485	2	9.4742	2.99	<.10	.091
Error	190.0383	60	3.1673			
Within participants						
Time	0.3251	2	0.1626	0.27	*ns*	.004
Time x Group	4.2712	4	1.0678	1.75	*ns*	.055
Error	73.3079	120	0.6109			
Total	286.8910	188				

ANOVA table in the ETT (higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	3.1669	2	1.5834	0.44	*ns*	.015
Error	214.7828	60	3.5797			
Within participants						
Time	89.8762	2	44.9381	59.05	<.01	.496
Time x Group	6.0602	4	1.5150	1.99	<.10	.062
Error	91.3293	120	0.7611			
Total	405.2153	188				

ANOVA table in the EWT (the whole essay, higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	2.0107	2	1.0054	2.22	*ns*	.069
Error	27.2207	60	0.4537			
Within participants						
Time	0.2049	2	0.1024	0.53	*ns*	.009
Time x Group	0.7483	4	0.1871	0.96	*ns*	.031
Error	23.3402	120	0.1945			
Total	53.5248	188				

ANOVA table in the EWT (the present-counterfactual conditional, higher proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	1.7072	2	0.8536	2.21	*ns*	.069
Error	23.1779	60	0.3863			
Within participants						
Time	3.7969	2	1.8984	16.97	<.01	.221
Time x Group	0.8114	4	0.2028	1.81	*ns*	.057
Error	13.4242	120	0.1119			
Total	42.9176	188				

ANOVA table in the untimed GJT (the conditionals, lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	76.5878	2	38.2939	2.13	*ns*	.054
Error	1349.4470	75	17.9926			
Within participants						
Time	96.5743	2	48.2871	21.07	<.01	.219
Time x Group	49.7440	4	12.4360	5.43	<.01	.126
Error	343.8099	150	2.2921			
Total	1916.1630	233				

ANOVA table in the untimed GJT (the present-counterfactual conditional, lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	30.5810	2	15.2905	4.05	<.05	.096
Error	283.0993	75	3.7747			
Within participants						
Time	8.6883	2	4.3441	4.61	<.05	.058
Time x Group	26.2351	4	6.5588	6.95	<.01	.156
Error	141.4967	150	0.9433			
Total	490.1002	233				

ANOVA table in the ETT (lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	30.6288	2	15.3144	5.11	<.01	.120
Error	224.7342	75	2.9965			
Within participants						
Time	84.3557	2	42.1778	51.66	<.01	.408
Time x Group	27.2424	4	6.8106	8.34	<.01	.182
Error	122.4790	150	0.8165			
Total	489.4402	233				

ANOVA table in the EWT (the whole essay, lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	8.7124	2	4.3562	3.87	<.05	.094
Error	84.4326	75	1.1258			
Within participants						
Time	0.4653	2	0.2327	0.33	*ns* .004	
Time x Group	5.4443	4	1.3611	1.92	*ns*	.049
Error	106.2209	150	0.7081			
Total	205.2754	233				

ANOVA table in the EWT (the present-counterfactual conditional, lower proficiency group)

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	0.0511	2	0.0255	0.07	*ns*	.002
Error	27.0729	75	0.3610			
Within participants						
Time	3.6364	2	1.8182	12.10	<.01	.139
Time x Group	0.8805	4	0.2201	1.46	*ns*	.038
Error	22.5471	150	0.1503			
Total	54.1881	233				

Appendix H: Examples of the ETT

日本語を見て（　　　）内に正しい英語を書きなさい。（１語とは限らない）
＊指示があるまで「訂正→」の欄には何も記入しないこと
（　　　）組　（　　　）番　発　・　標　氏名（　　　　　　　　　　　　）

	問　　題	語　句	Check！
1	デイビットはすでに札幌に引っ越している。 David（　　　　　　　　　　　　　） to Sapporo. 訂正→ David（　　　　　　　　　　　　　） to Sapporo.	引っ越す　move すでに　already	5
2	その製品には2001年より税金が含まれている。 The product（　　　　　　　　　） taxes since 2001. 訂正→ The product（　　　　　　　　　） taxes since 2001.	含む　include	2 & 6
3	私たちは2015年にハンバーガーショップを開店した。 We（　　　　　　　　　） a hamburger shop in 2015. 訂正→ We（　　　　　　　　　） a hamburger shop in 2015.	開店する　open	3
4	彼は以前４回コンサートに行ったことがある。 He（　　　　　　　　　） to a concert four times before. 訂正→ He（　　　　　　　　　） to a concert four times before.		5
5	トムは今朝からあのベンチに座っている。 Tom（　　　　　　　） on that chair since this morning. 訂正→ Tom（　　　　　　　） on that chair since this morning.		5
6	彼は先週トラを見たとき大声を上げた。 He（　　　　　　　） when he saw a tigar last week. 訂正→ He（　　　　　　　） when he saw a tigar last week.	大声を上げる　cry out	3
7	ケンジは最近十分寝ている。 Kenji（　　　　　　　） enough sleep lately. 訂正→ Kenji（　　　　　　　） enough sleep lately.		4
8	子供のころから彼は彼女を嫌っている。 He（　　　　　　　） her since they were children. 訂正→ He（　　　　　　　） her since they were children.	嫌う　hate	2 & 6
9	ヒロシはこないだの冬にアデレードへ行った。 Hiroshi（　　　　　　　） to Adelaide last winter. 訂正→ Hiroshi（　　　　　　　） to Adelaide last winter.		3
10	アスカは以前その会議に２度参加したことがある。 Asuka（　　　　　　　） the meeting twice before. 訂正→ Asuka（　　　　　　　） the meeting twice before.	参加する　join	5
11	ここ７日間湿っている。 It（　　　　　　　） humid for the past seven days. 訂正→ It（　　　　　　　） humid for the past seven days.		6
12	トムは今朝からクリスと話している。 Tom（　　　　　　　） with Chris since this morning. 訂正→ Tom（　　　　　　　） with Chris since this morning.		1 & 6
13	彼女は昨日母のために花を摘んだ。 She（　　　　　　　） some flowers for her mother yesterday. 訂正→ She（　　　　　　　） some flowers for her mother yesterday.	摘む　pick	3
14	ボブは昨日将棋をした。 Bob（　　　　　　　） Shogi yesterday. 訂正→ Bob（　　　　　　　） Shogi yesterday.	（将棋を）する　play	3
15	ユウカはこれまでに隣人たちを助けたことがある。 Yuka（　　　　　　　） the neighbours until now. 訂正→ Yuka（　　　　　　　） the neighbours until now.		5
16	彼は１０年前ピアニストだった。 He（　　　　　　　） a pianist ten years ago. 訂正→ He（　　　　　　　） a pianist ten years ago.		3
17	ケイトは先週の金曜からとても疲れている。 Kate（　　　　　　　） very tired since last Friday. 訂正→ Kate（　　　　　　　） very tired since last Friday.		6

Appendix I: Feedback sheet (as metalinguistic written CF)

現在完了形とは「**過去から現在まで続く状況や気持ち**」を表すものです。

■【過去形との違い】

> A:　Satoshi <u>lost</u> his wallet.　　　　　（過去形）
> B:　Satoshi <u>has lost</u> his wallet.　　　（現在完了形）

　Aは単に「（財布を）なくした」という過去の事実のみを表し，現在と関連はありません。よってまだなくしたままかもしれないし、見つかっているかもしれません。
　一方、Bは「なくした」という過去の事実だけでなく、<u>なくしたままであるという現在の状況</u>まで表します。

■【現在完了の形】
現在完了は〈have/has＋過去分詞〉で表します。
＊過去のある時から現在までの**動作の継続**を表すには，完了進行形を用います。　**Check 1**
　Bob has been cleaning his room since this morning.
「ボブは今朝からずっと自分の部屋の掃除をしています。」
　ただし，ある状態の継続を表す**状態動詞**は進行形にはなりません。　**Check 2**
　× I **have been belonging** to tennis club since I was the first year student.

■【現在完了形の意味】
１．完了：　　「〜した（ところだ）」
２．経験：　　「（今までに）〜したことがある」
３．継続：　　「（今まで）ずっと〜である」

□「状態動詞」の例
know（知っている）, remember（覚えている）, understand（理解している）, like（好んでいる）, hate（嫌っている）, have（持っている）, differ（異なっている）, resemble（似ている）, depend（依存している）, remain（〜のままである）, contain/include（〜を含んでいる）, owe（〜を支払う義務がある）等

重要！　**Check 3**

▽現在完了とともに使われない語句
yesterday（昨日）, last night（昨夜）, last week/month/year（先週 / 先月 / 昨年）, ago（〜前に）, just now（つい先ほど）, in 2001（〜年に）, when S＋V（SがVした時）等
　現在完了は現在の状況に重点を置いた表現なので，過去のある時点を表すような語句とともには用いられない。
　例）○　He left a few minutes **ago**.　　　　　×　He has left a few minutes **ago**.

▽現在完了とともに使われる語句
　【 完了 の意味で 】just（たった今）, already（もうすでに）,　lately（最近）, this week/month（今週 / 今月）
Check 4
　　　　　例）I have **already** finished lunch.
　【 経験 の意味で 】ever（「否定文・疑問文で」これまでに）, never（これまで一度も〜ない）, before（「漠然と」以前に）, once/twice/three times（「回数を表す」 1 / 2 / 3度）, until now（今までに）
Check 5
　　　　　例）I have been to Scotland **twice before**.
　【 継続 の意味で 】for（「期間」〜の間）for the past〜（ここ〜間）, since（〜以来）
Check 6

Appendix J: Examples of the tests

The timed grammaticality judgement test (timed GJT)

Slides

<table>
<tr>
<td>

Pretest

<div align="center">

英文〇×問題

リラックスして解いてください。
解答は**解答用紙**へ記入してください。

</div>
</td>
<td>
<div align="center">

答え方

</div>

英文に**誤りがない（正しい英文）**
　　　　　　　　　→　　**〇の欄に**✓
英文に**誤りがある（間違った英文）**
　　　　　　　　　→　　**×の欄に**✓
</td>
</tr>
<tr>
<td>
<div align="center">

答え方

制限時間があるので、
素早く判断し解答してください。

</div>
</td>
<td>

1

She cried out when she saw a ghost last week.
</td>
</tr>
<tr>
<td>

2

Masako lived in Okayama since last winter.
</td>
<td>

3

Miku has learned Spanish since 1980.
</td>
</tr>
</table>

Answer sheet for the timed GJT

英語〇×問題　解答用紙

解答方法：英文に誤りがない（正しい英文）→　〇の欄に✓

英文に誤りがある（間違った英文）→　×の欄に✓

		〇	×
練習問題	1		
	2		
	3		

		〇	×		〇	×
問題	1			14		
	2			15		
	3			16		
	4			17		
	5			18		
	6			19		
	7			20		
	8			21		
	9			22		
	10			23		
	11			24		
	12			25		
	13			26		

（　）組（　　）番　　発　・　標　　氏名（　　　　　　　　　　　）

The elicited imitation test (EIT)

Slides

Oral Elicited Information Test

暗示的 知識 測定 テスト

Pretest (very short ver.)

手 順

1．読まれる英文を聞く。

2．スクリーンに現れた日本語の質問に、
　　「はい・いいえ」で答える。

3．その後、最初に聞いた英文を思い出して言う。
　　（ただし**英文に誤りがある場合は正しく直した英文**を言うこと。）

第１問

1

音声を聞いてください。

1

あなたは英検やＴＯＥＦＬを受けて
みたいと思いますか？

指 示

聞いた英文を言ってください。
（間違いがあれば直して言うこと）

The untimed grammaticality judgement test (untimed GJT)

英文を読み，間違いがなければ○を，間違いがあれば×を記入しなさい。
なお，間違いがある場合は，その個所に下線を引き，その下に正しい英語を書きなさい。

問題No.	○/×	英　文
例	○	I enjoyed swimming in the pool yesterday.
例	×	I go to America last summer. went
1		I made an excuse to my teacher when I was late for the game.
2		I have failed the test yesterday.
3		He has run since this morning.
4		Ken became sad lately.
5		Miku went to Kyoto twice before.
6		Bob has worked at this company since last year.
7		I have gone to London last summer.
8		They have been remaining good friends since they joined the seminar.
9		David drank vodka until now.
10		Toshio has already closed his restaurant.
11		He has been born in 1976.
12		It was snowing for the past three days.
13		I went to New Zealand last fall.
14		Miku had stayed in this hotel three times before.
15		Ben already reported the result to his wife.
16		My opinion has been differing from yours since the discussion started.
17		He has often stayed at his grandmother's house when he was young.
18		I saw the car crash yesterday.
19		Toshio ran his own shop since he was in his 20s.
20		I have met Masako yesterday.
21		It has been hot for the past two days.
22		David has been interested in Jazz lately.
23		Toshio attended the conference three times before.
24		I have been a singer twenty years ago.
25		I was a cook four years ago.
26		The guys have moved the piano since 9 o'clock.
		（　　）組　（　　）番　発 ・ 標　氏名（　　　　　　　　　　　　　　）

153

Appendix K: ANOVA tables

ANOVA table in the timed GJT among higher proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	16.3183	2	8.1592	1.33	*ns*	.052
Error	299.4771	49	6.1118			
Within participants						
Time	30.1748	2	15.0874	7.80	<.01	.137
Time x Group	0.9908	4	0.2477	0.13	*ns*	.005
Error	189.5005	98	1.9337			
Total	536.4616	155				

ANOVA table in the untimed GJT among higher proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	112.7046	2	56.3523	1.83	*ns*	.070
Error	1505.4561	49	30.7236			
Within participants						
Time	76.1586	2	38.0793	18.09	<.01	.270
Time x Group	146.8347	4	36.7087	17.44	<.01	.416
Error	206.2456	98	2.1045			
Total	2047.3996	155				

ANOVA table in the timed GJT among lower proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	5.8447	2	2.9224	0.44	*ns*	.014
Error	402.1474	61	6.5926			
Within participants						
Time	17.6200	2	8.8100	2.93	<.10	.046
Time x Group	1.6643	4	0.4161	0.14	*ns*	.005
Error	366.6476	122	3.0053			
Total	793.9240	191				

ANOVA table in the untimed GJT among lower proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	93.0090	2	46.5045	1.79	*ns*	.055
Error	1585.3416	61	25.9892			
Within participants						
Time	124.0822	2	62.0411	26.45	<.01	.303
Time x Group	46.0691	4	11.5173	4.91	<.01	.139
Error	286.1521	122	2.3455			
Total	2134.6540	191				

ANOVA table in EIT among higher proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	0.4995	1	0.4995	0.09	*ns*	.003
Error	153.9143	27	5.7005			
Within participants						
Time	15.6072	1	15.6072	7.48	<.05	.217
Time x Group	0.1589	1	0.1589	0.08	*ns*	.003
Error	56.3238	27	2.0861			
Total	226.5038	57				

ANOVA table in EIT among lower proficiency learners

Source	SS	df	MS	F	P	Effect size (η_p^2)
Between participants						
Group	0.3542	1	0.3542	0.12	*ns*	.005
Error	74.7940	25	2.9918			
Within participants						
Time	14.4616	1	14.4616	24.00	<.01	.500
Time x Group	0.0172	1	0.0172	0.03	*ns*	.001
Error	14.4643	25	0.5786			
Total	104.0913	53				

Appendix L: The ETT
 Pretest for present perfect

英作文問題
N o . 1

ヒント

1) あなたはもう岡山に到着しましたか。

_____

_____

2) サムは毎朝７時に起きないのですか。はい、起きません。

_____

_____

3) 私は飛行機で旅行したことがない。

飛行機　airplane

_____

_____

4) 彼は宿題を終えることができたので、外出した。

_____

_____

5) 私はちょうど宿題を終えたところだ。

_____

_____

6) 以前この辺りに古い教会があった。

教会 church

_____

_____

7) 私は３年間ずっと英語を勉強している。

_____

_____

Score

（　　　）組　（　　　）番　発　・　標　氏名（　　　　　　　　　）

157

Posttest for present perfect

英作文問題

ヒント

1) 私はまだ昼食を終えていない。　　　　　　　　　　　　　　　　　　　昼食　lunch

_____

_____

2) ベンはテニス部員ではないのですか。はい、部員ではありません。　　部員 member

_____

_____

3) 私は７時からずっとこのテレビ番組を見ている。　　　　　　　　　　テレビ番組
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　TV program

_____

_____

4) 私たちは先週ついにあの山を登ることができた。

_____

_____

5) 私たちは子供のころからお互いを知っている。　　　　　　　　　　　お互いeach other

_____

_____

6) 以前ここに三つホテルがあった。

_____

_____

7) あなたはこれまでに京都に住んだことがありますか。

_____

_____

Score

（　　）組　（　　）番　発 ・ 標　氏名（　　　　　　　　）

158

Delayed posttest for present perfect

英作文問題

ヒント

1) 私は先週からずっと忙しい。

_____

_____

2) ケンは毎日部活はしないのですか。はい、しません。

部活
club activity

_____

_____

3) 私は彼の家に訪れたことがない。

_____

_____

4) ついに彼は昨日お気に入りの俳優に会うことができた。

俳優　actor

_____

_____

5) あなたはもう彼女に電話しましたか。

_____

_____

6) 以前この辺りに郵便局があった。

郵便局　post office

_____

_____

7) 私はあなたを２時間ずっと待っている。

_____

_____

Score

（　　）組　（　　）番　発　・　標　氏名（　　　　　　　　　　）

Pretest for past perfect

英作文問題

ヒント

1) 私は３０歳になるまで，イギリスに訪れたことはなかった。

_____

_____

2) あなたは忙しくありませんよね。はい、忙しくありません。

_____

_____

3) 彼が現れるまで，私は３時間ずっと待っていた。

現れる
appear

_____

_____

4) 先生が来たら私は英語の勉強を始めよう。

_____

_____

5) 私たちはあの授業が休講になっていたこと（キャンセルされていたこと）を知らなかった。

キャンセルする
cancel

_____

_____

6) トムがいつ日本を発つか知らない。

_____

_____

7) あなたが彼女に電話したとき，彼女は家を出てしまっていましたか。

_____

_____

Score

（　　）組　（　　）番　発 ・ 標　氏名（　　　　　　　　　　　）

Posttest for past perfect

英作文問題　　　　　　　　　　　No. 2

1)　あなたが現れるまで，私はずっと手紙を書いていた。

現れる　appear

_____

_____

2)　あなたは約束を破りませんよね。はい、破りません。

約束を破る
break one's promise

_____

_____

3)　彼女は自分の写真が撮られていたとは思わなかった。

撮る　take

_____

_____

4)　暗くなる前に私たちは家に着くだろう。

暗くなる
get dark

_____

_____

5)　あなたが家に帰ったとき，母親はすでに夕食を作り終えていましたか。

_____

_____

6)　ケンが明日私のところに訪れるかどうか知らない。

_____

_____

7)　日本に来るまで，私は納豆を食べたことはなかった。

納豆　natto

_____

_____

Score

（　　　）組　（　　　）番　発　・　標　　氏名（　　　　　　　　　　　）

Delayed posttest for past perfect

英作文問題 No.3

ヒント

1) 私は自分の名前が呼ばれていたことに気づかなかった。

_____

_____

2) あなたは喉が渇いていないですよね。はい、渇いていません。

のどが渇いている
thirsty

_____

_____

3) あなたが彼の家に行ったとき，彼は宿題を済ませてしまっていましたか。

_____

_____

4) 彼が出て行ったらケーキを食べよう。

_____

_____

5) 高校生になるまで，飛行機を利用したことはなかった。

飛行機　plane

_____

_____

6) ジャネットがいつ東京に引っ越すか知らない。

_____

_____

7) 彼女が現れるまで，ずっと音楽を聴いていた。

現れる　appear

_____

_____

Score

（　　）組　（　　）番　発　・　標　　氏名（　　　　　　　　　　）

162

Appendix M: Feedback sheet (as metalinguistic written CF)

１．過去完了形

過去完了には２つの用法があります。

①過去の時点を基準にして現在完了と同じ「完了・結果，経験，継続」の意味を表す用法

②過去の時点よりも過去のことを表す用法

【①過去の時点を基準にして現在完了と同じ「完了・結果，経験，継続」の意味を表す用法について】

以下の２つの英文を見てください。

A: I <u>have</u> been busy for a week.　（現在完了）

B: I <u>had</u> been busy for a week.　　（過去完了）

Ａは現在完了形が用いられており、１週間前から今現在まで１週間忙しかったという意味になります。一方、Ｂは過去のある時点からある時点までの１週間忙しかったという意味になります（今現在忙しいかどうかは不明です）。

よってＢのような過去完了の文では、過去のある時点を明確に表す表現とともに使われるのが普通です。以下の英文であれば、"When I reached the station,"の部分が過去の時点を表しています。

<u>When I reached the station</u>, the train had already gone.

　訳：　「私が駅に着いたとき、電車はすでに出てしまっていた」

過去完了は〈had + 過去分詞〉で表します。

否定形：　had + not/never + 過去分詞

疑問形：　Had + 主語 + 過去分詞〜？

＊過去のある時からある時までの<u>動作の継続</u>を表すには，過去完了進行形を用います。

　I had been waiting for two hours when he appeared.

　訳：　「私は彼が現れるまで２時間ずっと待っていた」

【過去完了形の意味】

（１）完了・結果：　「〜してしまっていた（〜した）」

（２）経験：　「（ある過去の時点までに）〜したことがあった」

（３）継続：　「（ある過去の時点まで）ずっと〜していた」

【②過去の時点よりも過去のことを表す用法について】

以下の２つの英文を見てください。

   I lost the bag that I had bought a week before. （過去完了）

  訳： 「私は１週間前に買ったバッグをなくした」

   I didn't know that the window had been broken. （過去完了＋受動態）

  訳： 「私はその窓が壊されたのを知らなかった」

①の過去完了は現在完了を過去の時点にずらしたものですが、現在完了とは関係なく、単に過去の時点よりもさらに過去の出来事を表すこともできます。

２．否定形の疑問文に対する答え方

英語は疑問文が否定形であってもなくても、返事の内容が肯定ならば Yes,否定ならば No で答えます。日本語のはい・いいえに惑わされないように！

以下の英文を見てください。

  Mr. Aoyama isn't a mathmatics teacher, is he? No, he isn't.

  訳： 「青山先生は数学の先生ではありませんよね。」「はい、数学の先生ではありません。」

このように付加疑問文の場合も否定形の疑問文ですので同じルールが当てはまります。
青山先生は英語の先生であり、数学の先生ではないので返事の内容は No となります。日本語の「はい」に惑わされないようにしましょう。

３．副詞節と名詞節の未来表現

時や条件を表す副詞節内では、内容が未来のことを表していても現在形を使用します。一方、名詞節内では、未来の内容は未来表現を使用します。

以下の英文を見てください。

  We will arrive at the gallary before it rains.

  訳： 「雨が降る前に私たちはギャラリーに着くだろう。」

日本語訳をみるとまだ雨は降っていないので未来表現で表すと思ってしまいますが、before it rains は副詞節なので未来の内容ですが現在形を用います。

  I don't know if it will rain.

  訳： 「雨が降るかどうか分からない」

一方、上の例文では if it will rain の部分は know の目的語であり名詞節なのでは未来の内容であるので未来形を用います。

Appendix N: ANOVA tables

ANOVA table in higher item-specific proficiency group

Source	SS	df	MS	F	P	Effect size (η^2)
Between subjects						
Group	12.6934	3	4.2311	1.55	*ns*	.024
Error	235.4676	86	2.7380			
Within Subjects						
Time	9.6717	2	4.8358	3.21	<.05	.018
Time x Group	10.0919	6	1.6820	1.11	*ns*	.019
Error	259.4907	172	1.5087			
Total	527.4152	269				

ANOVA table in higher item-specific proficiency group (only pretest to posttest)

Source	SS	df	MS	F	P	Effect size (η^2)
Between subjects						
Group	6.0913	3	2.0304	1.24	*ns*	.024
Error	141.0208	86	1.6398			
Within Subjects						
Time	7.0646	1	7.0646	6.71	<.05	.028
Time x Group	9.2324	3	3.0775	2.92	<.05	.036
Error	90.5764	86	1.0532			
Total	253.9856	179				

ANOVA table in middle item-specific proficiency group

Source	SS	df	MS	F	P	Effect size (η^2)
Between subjects						
Group	43.4940	3	14.4980	1.24	*ns*	.029
Error	608.6905	52	11.7056			
Within Subjects						
Time	67.5833	2	33.7917	4.61	<.05	.044
Time x Group	37.4167	6	6.2361	0.85	*ns*	.025
Error	761.6667	104	7.3237			
Total	1518.8512	167				

ANOVA table in middle item-specific proficiency group (only pretest to posttest)

Source	SS	df	MS	F	P	Effect size (η^2)
Between subjects						
Group	27.2411	3	9.0804	0.98	*ns*	.032
Error	481.2500	52	9.2548			
Within Subjects						
Time	58.5804	1	58.5804	12.49	<.01	.069
Time x Group	35.9554	3	11.9851	2.55	<.10	.042
Error	243.9643	52	4.6916			
Total	846.9911	111				

ANOVA table in lower item-specific proficiency group

Source	SS	df	MS	F	P	Effect size (η^2)
Between subjects						
Group	76.1333	2	38.0667	4.22	<.05	.052
Error	513.6667	57	9.0117			
Within Subjects						
Time	310.9000	2	155.4500	43.18	<.01	.214
Time x Group	142.6667	4	35.6667	9.91	<.01	.098
Error	410.4333	114	3.6003			
Total	1453.8000	179				

Appendix O: Questionnaire

<div align="center">

英作文上の誤りに与えられる訂正に関するアンケート

ご協力のお願い

</div>

このアンケートは，英作文上に現れる誤りに対して与えられる訂正に対するみなさんの意見をうかがうことを目的としています。質問を読み，最も当てはまる番号ひとつに〇をつけ，その理由を空所に書いてください。回答しづらい質問があるかもしれませんが，あまり考えこまず答えていただいて構いません。よろしくお願いします。

<div align="right">3年英語科　青山</div>

■3年（　　　）組（　　　）番　発・標　名前（　　　　　　　　　　　　　　）

Q1：英作文の誤りは誰に訂正してもらいたいですか。

　　　1．先生　　　　2．友達（同級生）　　　3．どちらでもよい

```
┌─────────────────────────────────────────┐
│                                         │
│                                         │
│                                         │
└─────────────────────────────────────────┘
```

Q2：誤りはどのように対処してもらいたいですか。

　　　1．最初から正しい答えを教えてもらう

　　　2．最初はヒントだけもらい自分で訂正してから、最後に正しい答えを教えてもらう

　　　3．どちらでもよい

```
┌─────────────────────────────────────────┐
│                                         │
│                                         │
│                                         │
└─────────────────────────────────────────┘
```

Q3：誤りに対する訂正はいくつしてもらいたいですか。

　　　1．すべての誤りを訂正してもらいたい

　　　2．特定の文法項目ひとつに絞って訂正してもらいたい（例：冠詞なら冠詞のみ）

　　　3．どちらでもよい

```
┌─────────────────────────────────────────┐
│                                         │
│                                         │
│                                         │
└─────────────────────────────────────────┘
```

Q4：訂正された英作文が返却された後は，たいていどのようなことをしますか。

　　　1．訂正をもとに，なぜ間違えたかを考え，書き直しをしている

　　　2．訂正をもとに，なぜ間違えたかを考えるが，書き直しはしていない

　　　3．見直すことはない（特になにもしない）

```
┌─────────────────────────────────────────┐
│                                         │
│                                         │
│                                         │
└─────────────────────────────────────────┘
```