

Corrective feedback in computer-mediated collaborative writing and revision contributions

Taichi Yamashita, Iowa State University

Abstract

This study investigated the effects of corrective feedback (CF) during in-class computer-mediated collaborative writing on grammatical accuracy in a new piece of individual writing. Forty-eight ESL students at an American university worked on two computer-mediated animation description tasks in pairs. The experimental group received indirect CF on English indefinite and definite articles from the researcher during the tasks, while the comparison group worked on the same tasks without CF. Each computer screen was recorded during the treatment, so that the number of revision contributions from each individual learner could be identified. L2 development was measured by a pretest, posttest, and delayed posttest, where the students worked on an animation-description task without a partner. A repeated-measures ANOVA indicated a significant relationship between the presence of CF and accuracy improvement over time. Furthermore, multiple regression analyses suggested a significant relationship between the number of learners' revision contributions and the delayed posttest scores when the pretest scores held constant. That is, individual learners' long-term L2 development varied depending on the extent to which they contributed to the revision. These findings demonstrate the importance of tracking individuals' contributions while calling for more detailed collection of data on actual revisions and the distribution of revision work within pairs or groups.

Keywords: *Computer-Mediated Communication, Second Language Acquisition, Writing, Collaborative Learning*

Language(s) Learned in This Study: *English*

APA Citation: Yamashita, T. (2021). Corrective feedback in computer-mediated collaborative writing and revision contributions. *Language Learning & Technology*, 25(2), 75–93. <http://hdl.handle.net/10125/73434>

Introduction

In the past decades, many studies have explored collaborative writing, “an activity where there is a shared and negotiated decision making process and a shared responsibility for the production of a single text” (Storch, 2013, p. 3). This sort of task allows learners to experience a sense of co-authorship at multiple stages of writing and thus engenders interactions which are expected to facilitate their L2 development. The current body of research has identified a variety of factors influencing task processes and L2 development, such as task type, method of grouping or pairing learners, and interaction patterns (see Storch, 2013 for review).

However, Ammar and Hassan (2018) indicated that very limited attention has been paid to the role of instructors while learners are working on a collaborative writing task. Accordingly, relatively little is known about how and to what extent instructors' assistance facilitates learners' subsequent performance. In particular, studies have been inconclusive as to the extent to which learners can develop their grammatical accuracy, one of the critical aspects of writing for comprehensibility. Therefore, the present study investigates the effects of corrective feedback (CF)—one form of instructors' efforts to draw learners' attention to grammatical issues—provided by harnessing the synchronous interface of Google Docs™.

Furthermore, the study addresses one of the most commonly raised concerns regarding collaborative tasks, uneven work distribution, by examining to what extent individual learners' shares of contribution to revision in response to instructors' CF predict their L2 development.

Literature Review

The Effects of Collaborative Writing on Individuals' L2 Development

According to Storch (2013), the benefits of collaborative writing for SLA are theoretically supported from a cognitive perspective. Drawing on the Interaction Hypothesis, she claims that collaborative writing requires multiple students to be responsible for a single text, and this co-authorship creates an acquisition-rich context by affording them opportunities to negotiate meaning and form (Long, 2007). Furthermore, the slow pace and permanent nature of written production are considered to effectively draw learners' attention to language form, pushing them to extend their current linguistic knowledge (Swain, 2006). Composing a single text in a pair or group is also common practice in content classes. Thus, ESL learners, especially those who need to achieve academic success in English (e.g., English for Academic Purposes), are expected to deal with interactional task processes, such as suggesting ideas and responding to peers' suggestions, after they meet their language requirement. Therefore, collaborative writing tasks in ESL classes have the potential to prepare students for real-world tasks, as well as supporting their SLA.

Collaborative writing research appears to have mostly adopted a process-oriented perspective which emphasizes L2 development in action. In particular, many collaborative writing studies have described learners' dialogue during a task (mostly referred to as language-related episodes, LRE), operationalizing the dialogue as a learning opportunity (e.g., Wigglesworth & Storch, 2009). The predominance of this methodological orientation seems to have marginalized inquiry as to the extent to which learners improve their own performance (when working without relying on external sources, e.g., group members, teacher) as a result of collaborative writing (Storch, 2013). In this regard, collaborative writing may need to be evaluated in terms of its impact on L2 development in line with expected learning outcomes laid out in a syllabus, such as grammatical accuracy and lexical choice, by assessing learners' independent performance.

A few studies, adopting a product-oriented approach, have investigated individuals' grammatical development by employing the pretest and posttest design. In paper-based collaborative writing, it seems that, although collaborative writing engenders learning opportunities (e.g., Kuiken & Vedder, 2002; Reinders, 2009; and many others), learners working on writing tasks with peers do not necessarily outperform their counterparts who work on the same tasks individually (Kuiken & Vedder, 2002; Reinders, 2009; Shehadeh, 2011). In a computer-mediated platform, Bikowski and Vithanage (2016) had 59 ESL college students work on four collaborative writing tasks either individually or in a group. Results from a pretest and posttest showed a significant difference between the groups. However, the use of only overall scores, which included content, organization, academic style, and grammar, leaves it unclear as to what extent the difference can be ascribed to changes in their grammatical accuracy. Thus, the current body of research does not provide clear evidence regarding the relative benefits to students' grammatical development of working on writing tasks with peers versus working on them individually.

One of the potential factors accountable for these inconclusive findings is the tendency that learners discuss lexical items rather than grammatical items during collaborative writing tasks, as has been revealed by studies which examined task processes during collaborative writing (Coyle & Roca de Larios, 2014; Kim & McDonough, 2011; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). For instance, Coyle and Roca de Larios (2014), who had primary school pupils work on a narrative writing task in pairs, found that there were almost no grammatical features noticed during the composition. Similarly, Wigglesworth and Storch (2009) examined the collaborative dialogue of college students during an argumentative writing task, finding that more than half of the dialogue addressed lexical items, with only one third targeting form-

related features. These findings seem to suggest the need of instructors' intervention to draw learners' attention to grammatical features when appropriate.

Teacher-provided CF in Collaborative Writing

Learners' attention can be drawn with various techniques, and the present study particularly focuses on corrective feedback (CF), the role of which has been one of the long-standing topics in second language writing research (Riaz et al., 2018). Although a series of publications by Truscott has set out a critical viewpoint on written CF (e.g., Truscott, 1996, 1999, 2007), the current body of research appears to reach a conclusion that written CF is at least effective for improving accuracy of certain grammatical items under certain conditions, as far as individual writing tasks are concerned (Bitchener & Storch, 2016). In fact, a meta-analysis carried out by Kang and Han (2015) showed that written CF leads to improvement, with moderate effect sizes. However, few studies have investigated the effects of CF in collaborative writing (Wigglesworth & Storch, 2012b).

Previous studies have mostly addressed how a pair of learners deals with CF provided after task completion, reporting that learners notice and incorporate some forms into their subsequent rewriting. Swain and Lapkin (2002), for example, investigated two French immersion students' interactions while comparing their original text with a reformulated version of their text. The two learners incorporated many linguistic items from the reformulated version when directed to work on the same writing prompt individually again. Coyle and Roca de Larios (2014) compared the effects of providing model texts or direct CF, to 46 L2 learners of English in a Spanish primary school. They found that direct CF led to significantly greater accuracy in a collaboratively rewritten text than access to model texts. Another two studies examined the effects of providing CF in comparison with the effect of providing no feedback or model text. Adams (2003) recruited 56 L2 learners of Spanish and found that those in the experimental groups, who received reformulation (i.e., the rephrased form of non-target-like expressions), produced target-like forms in their individual rewriting significantly more than learners in the comparison group. Wigglesworth and Storch (2012a) investigated effects of providing reformulation, editing symbols (i.e., the indication of the type of each error), and no CF, with 72 ESL learners. The results indicated that reformulation led to greater accuracy in the collaborative rewriting than provision of editing symbols and provision of no CF. Thus, these studies on the effects of CF given after task completion imply the positive effects of CF during collaborative writing.

However, studies of collaborative writing face two primary methodological problems. First, the learners' dialogue during collaborative writing has often been situated in a laboratory setting. While this study design ensures collection of detailed and accurate data on interaction, such a controlled setting may pose a question regarding the ecological validity. For instance, in the study by Swain and Lapkin (2002), though the participants were reportedly 'quite at ease' and 'not intimidated' (p. 290), their observation by three research assistants may have led to interactions which the participants would not demonstrate in the classroom setting. This limitation can be found in Adams (2003) and Wigglesworth and Storch (2012a), where one pair at a time worked on a task under an administrator's control. In fact, classroom research has shown that pair work does not always turn out to be collaborative, and the frequency of interactional features varies considerably (McDonough, 2004; Storch, 2001, 2002). Another limitation is that students' learning has been measured by having them work on the same writing task, with the same prompt, after the CF treatment. This threatens the validity of any conclusions regarding learners' mental representations of linguistic knowledge. That is, unless their development is measured in a new piece of writing, it remains questionable as to whether they have understood the rules behind linguistic manifestations in the revision phase or simply reused some memorized chunks (Truscott, 1996). Therefore, it remains unknown to what extent learners can apply what they learn from teacher-provided CF in a collaborative writing task to a novel writing task without a partner (Storch, 2013).

Uneven Revision Contribution and L2 Development

One of the most commonly raised concerns regarding collaborative writing is uneven work distribution, because individual learners not only share but also diffuse the sense of the ownership of a single text (Storch, 2013). Work distribution appears in various forms, such as the number of words written by each group member (Zhang, 2019). Another form of work distribution appears in the process of revision. Specifically, collaborative writing does not ensure that individual learners have opportunities to respond to teacher-led CF because they share not only a text to be completed but also a text to be corrected, and the opportunities to revise are especially reduced when there are limited possibilities of revision of the linguistic errors (e.g., in the case of morphosyntactic errors). For instance, when a pair of learners receives CF, one learner may revise the error without allowing her partner the opportunity to work out the required revision. The management of revision processes is a variable of great relevance to both written CF and collaborative writing studies. Written CF research is inconsistent as to whether learners are asked to revise their errors (e.g., Benson & DeKeyser, 2019; Karim & Nassaji, 2020) or only to look at the given CF for a very short time period (e.g., Bitchener & Knoch, 2008; Sheen, 2007; Stefanou & Révész, 2015). Meanwhile, collaborative writing research seems to have kept to a common methodological approach which only examines revisions per pair without tracking who produced each revision (e.g., Adams, 2003; Coyle & Roca de Larios, 2014; Wigglesworth & Storch, 2012a).

These methodological conventions appear to render it challenging to draw a solid conclusion on the effects of written CF from a cognitive perspective. Two considerations are important here. Firstly, revision can indicate learners' noticing and form-meaning mapping (Schmidt, 1990). Thus, individual learners who produce more revisions are expected to demonstrate L2 development but at the same time to diminish their fellow group members' learning opportunities. Secondly, skill acquisition theory indicates that learners who make revisions benefit from proceduralizing the declarative knowledge, enabling them to produce features more accurately, quickly, and effortlessly (DeKeyser, 2015). Furthermore, transfer-appropriate processing suggests that learners who produce the correct form in response to CF (e.g., the indefinite article) to convey certain meaning (e.g., unknown to the addressee) may be able to produce the linguistic feature more accurately in a novel, meaningful context than those who do not (Lightbown, 2008). In short, learners who make revisions are more likely to improve their performance in a new piece of writing which is relatively meaning-oriented.

Yet very few studies have empirically isolated the effects of revision. Shintani et al. (2014) examined the effects of revision on accuracy in a new piece of writing. They had 171 college students work individually on a dictogloss task with the English indefinite article hypothetical conditional structure being targeted. The treatment of their four experimental groups differed in terms of CF type (i.e., metalinguistic explanation vs. direct CF) and whether the learners were allowed to make revisions upon their receipt of CF. The results indicated that, while the revision opportunity did not appear to be associated with any significant difference on an immediate posttest, the two groups with the revision opportunity significantly outperformed the comparison group on a delayed posttest, suggesting its impact on long-term learning. In contrast, the two groups without such revision opportunity did not outperform the comparison group. Based on this result, the researchers speculated that learners process written CF more deeply and consolidate their declarative knowledge by revising their errors.

Due to the scarcity of research, more studies are needed in order to better understand the relationship between individuals' revision contributions and L2 development. In particular, if the aforementioned speculated moderating role of the revision contribution is empirically supported, this will suggest that revision per pair in collaborative writing does not necessarily indicate individuals' L2 development, but potentially overestimates the learning of individuals who rarely contribute to revision. Furthermore, if revising is related to L2 development, either negatively or positively, it would be safe to argue that findings from written CF studies are possibly confounded by the number of revisions individual learners make. This would suggest the need for a more transparent data report than has been provided thus far in some studies,

to show not only whether a study invites learners to revise their errors (Liu & Brown, 2015) but also whether learners *actually* revise errors. In short, examining the relationship between individuals' revision contributions in collaborative writing and their L2 development is expected to further raise methodological and reporting awareness of researchers in written CF and collaborative writing, as well as instructors' awareness of the need to track individuals' contributions to a writing task instead of estimating learning per pair.

Direct and Indirect CF

Ellis (2009) has provided an informative typology of written CF. One of the long-standing debates addresses the comparative effects of direct and indirect CF (e.g., Van Beuningen et al., 2012). Direct CF refers to the provision of the correct form, whereas indirect CF refers to the indication of the presence of an error (e.g., by underlining or circling) while withholding the correct form. It has been claimed that direct CF facilitates L2 development because it provides both positive and negative evidence at once, whereas indirect CF is more likely to result in learners' revision engagement and thus in longer-term retention (Chandler, 2003; Ferris & Roberts, 2001). The current body of research suggests overall that direct CF is more effective than indirect CF when learners' grammatical knowledge is assessed immediately after they receive written CF treatment (Kang & Han, 2015).

The relationship between revision contribution and L2 development may be further mediated by this factor of CF type. For example, when a pair of learners receives direct CF, both learners may benefit regardless of their revision contribution because they are both exposed to the correct form. In contrast, use of indirect CF would possibly make this relationship more salient, because learners' responses to indirect CF should constitute a better indication of their noticing than their responses to direct CF, given the need to work out their error by themselves. Due to the exploratory nature of its inquiry, the present study only examines indirect CF, on the grounds of an expectation that L2 development as a result of indirect CF in collaborative writing is dependent on how individual learners act upon it.

Research Questions

Informed by the literature review above, the present classroom research aims to serve as a first attempt to investigate the effects of teacher-led CF during collaborative writing on accuracy in a new piece of individual writing, and to examine whether the accuracy is related to the individual's revision contribution. The following research questions were formulated:

1. To what extent is teacher-provided CF during computer-mediated collaborative writing effective for improving the accuracy of use of English definite and indefinite articles in a new piece of individual writing?
2. To what extent does the number of learners' individual revision contributions predict their short-term and long-term (i.e., two weeks later) learning regarding use of the articles?

Method

Participants

The present study recruited seven ESL writing classes at a large Midwestern university in the United States via convenience sampling. Learners enrolled in these classes were fully matriculated at the institution. They were placed into these classes after they failed to pass an in-house English placement test which assessed students' academic writing ability (e.g., organization, arguments, grammar, writing conventions). Of the recruited seven classes, two were a taking the first level in a sequence of two ESL courses, and the other five were taking the higher level course. In most cases, learners in the lower level course subsequently take the higher level course as a required course, and those completing the higher level course meet their language requirement and are not then required to take further ESL courses. All the classes met for three

hours per week for 16 weeks. Both courses aimed to prepare students for academic work at the institution by teaching a range of writing skills, such as how to write a thesis statement and how to avoid plagiarism. While sentence-level issues were emphasized more in the lower level course, with supplementary material provided on grammar, both courses included grammatical accuracy in their intended learning outcomes.

Fifty-two students agreed to participate, completing the informed consent form distributed prior to the data collection. They were randomly paired up within their class group (i.e., with each member of a pair at the same placement level), and each pair was randomly assigned to either an experimental ($n = 28$) or comparison group ($n = 24$), to minimize the moderating effects of proficiency and class membership. Four students missed one of the procedures and were therefore removed from the final data pool.

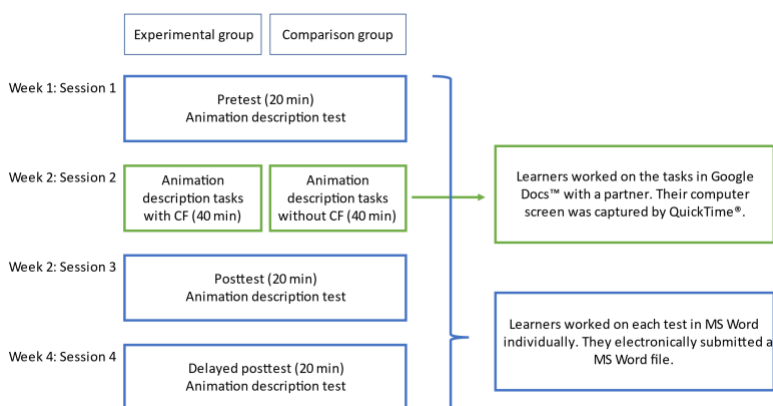
The final data pool had 26 students in the experimental group and 22 students in the comparison group. There were 19 females and 29 males whose age ranged from 18 to 29. The mean TOEFL iBT score, for those who had one, was 85.2 with a standard deviation of 10.19 ($n = 35$). Of those whose TOEFL iBT score was not available, eight reported their IELTS ($n = 5$: range 6.0–7.5), TOEFL PBT ($n = 2$: 540, 533), or TOEIC scores ($n = 1$: 655). No scores were reported by the remaining five. These test scores collectively indicate that the proficiency of the sample was approximately upper-intermediate to advanced. As far as their L1 was concerned, the largest group was of L1 Chinese speakers ($n = 20$), and the rest were L1 speakers of various languages, such as Arabic, Indonesian, and Japanese.

Research Procedures

The participants took a pretest one week before the treatment. On the treatment day, they were invited to a computer classroom, which was reserved for this research, and worked on the animation description tasks with their computer screen activity captured by QuickTime®. Two days after the treatment session, they took the posttest and the short survey. Lastly, they worked on the delayed posttest two weeks after the treatment (Figure 1). It should be noted again that learners worked on the treatment tasks in pairs in Google Docs™, whereas they completed each of the pretest, posttest, and delayed posttest without a partner, in Microsoft Word. All the procedures were implemented in their regular class hours.

Figure 1

Treatment and Test Procedures



Writing Task for Experimental and Comparison Groups

The task was Google Docs™-mediated collaborative writing, where the participants worked on animation description tasks in their designated pairs in their regular class hours. Google Docs™ was often adopted in each class, and thus it is safe to assume that the learners were familiar with the platform. In order to facilitate

face-to-face oral interactions, the paired learners were seated next to each other. They were given 20 minutes to describe an animation lasting approximately three minutes with their partner. During the writing task, one computer was available to each individual, and learners used their own institution account to work on their shared text in Google Docs™. Their computer screen contents were captured by QuickTime®. The animation was a *Tom and Jerry* cartoon. While previous written CF research has primarily employed picture description tasks (Liu & Brown, 2015), the animation description task was expected to be more successful in eliciting the target features simply because of the larger number of repeated and new objects that is possible in an animation storyline. Furthermore, animation description tasks differ from picture description tasks in that learners cannot see the whole story at a glance. Therefore, the animation description task was expected to be more engaging than picture description tasks, requiring the learners to play, stop, and replay scenes.

The learners watched two video clips which were comparable in length and thus were expected to take a comparable amount of time on task. In one video clip, Tom chases not only Jerry but also a kitten, who betrays Tom at the very beginning of the video. In the other video clip, Jerry plays a prank on Tom, asserting that Tom suffers from a serious disease and trying to treat him. These two clips were presented in a counterbalanced order to minimize any task order effects. The two clips were presented to half the pairs in one order and the other pairs in reverse order. Each pair was held accountable for the way they worked on the task, and they were encouraged to talk to each other while working. The entire writing task constituted a class hour of approximately 40 minutes. Each learner worked on both video clips with the same partner, and each computer screen was recorded during the writing task.

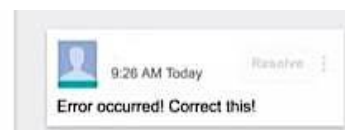
Feedback Operationalization

The present study adopted focused indirect CF (i.e., the indication of an error occurrence only for predetermined error types) (Ellis, 2009). The rationale for focusing on a single error type is that the study aimed to examine relationships between revision and subsequent performance in a new piece of writing. Indirect CF was chosen in an attempt to detect a relationship between the revision and L2 development, given that a learner's response to indirect CF should constitute a good indicator of their noticing because it withholds the correct form. The researcher shared one document with each pair through Google Docs™ and provided CF as soon as he detected an error, by highlighting the error with a comment indicating the error occurrence and prompting the learners to revise it (Figure 2).

Figure 2

Feedback Operationalization

Jerry throws the thermometer into Tom's mouth and use a lighter to heat the temperature up to proved Tom have a cold. After reading the Dr Quick's Treatment, Jerry let Tom to sit in a freeze



When learners incorrectly responded to CF, by, for example, simply omitting a misused definite article where the indefinite article was expected, or dismissing the CF by clicking on *Resolve* without correcting the error, CF was repeatedly provided in the manner shown in Figure 2. In the interface of Google Docs™, both students in a pair could see the CF provided by the researcher in a synchronous manner. Depending on the number of learners in a classroom, the researcher provided CF to three to four pairs during the treatment. During the treatment, there was no time solely dedicated to revision; instead, the learners were expected to respond to the CF by revising their errors when it suited them. After the treatment, classroom instructors were advised not to provide any explicit instruction on English articles until the delayed posttest was implemented.

Target Features

The target features were English indefinite and definite articles, and no CF was provided on other features during the treatment. In particular, the study focused only on the functions of [+Specific Referent, -Assumed Known to the Hearer] for the indefinite article and [+Specific Referent, +Assumed Known to the Hearer] for the definite article (Huebner, 1983). Simply put, if a countable noun refers to a specific object that is introduced into a narrative for the first time, it is accompanied by the indefinite article. On the other hand, a noun comes with the definite article when it refers back to an item that has been previously mentioned, regardless of its countability. L2 learners seem to acquire the definite article before the indefinite article and to acquire the indefinite article after demonstrating overgeneralized use of the definite article in [+Specific Referent, -Assumed Known to the Hearer] contexts (Chaudron & Parker, 1990; Master, 1997; Zdorenko & Paradis, 2008, 2012).

These grammatical items were selected for three reasons. Firstly, L2 learners, even those who are proficient enough to be fully matriculated at a U.S. college, appear to have difficulty in choosing the right article (Butler, 2002). This difficulty while acquiring the articles can be ascribed to, among other factors, their limited perceptual saliency (i.e., how easy it is to hear a structure) and, more likely, the challenge of having multiple semantic concepts to encode at one time (e.g., countable, specific, first mention) (DeKeyser, 1998; Ellis, 1990; Long, 2007). In fact, the sample of learners in this study also produced several instances of the overgeneralized definite article, suggesting their linguistic needs in the area of marking new and old information. Secondly, the choice of article is expected to contribute to textual coherence, especially in a descriptive type of writing, and thus errors in article use may result in confusion on the part of the reader. Additionally, the articles have been intensively investigated in written CF research (e.g., Bitchener, 2008; Bitchener & Knoch, 2008; Sheen et al., 2009) and collaborative writing research (e.g., Storch, 1999, 2007; Storch & Wigglesworth, 2010). This comparability with past studies was desired, especially because the present study was to examine an underexplored variable, the revision contribution.

Test Instruments

The study employed a pretest, posttest, and a delayed posttest. In each test, the participants were asked to watch an approximately three minutes long *Tom and Jerry* animation. Then they had 20 minutes to individually describe the clip in as much detail as possible using Microsoft Word. Three video clips, which were comparable in length to the two clips presented in the treatment, but different in content, were prepared, and the order of clips was varied to minimize test order effects. The three test video clips are labeled as A, B, and C in [Table 1](#).

Table 1

Counterbalancing of Tests

	Experimental (<i>n</i> = 26)			Comparison (<i>n</i> = 22)		
	6	10	10	8	9	5
Pretest	A	C	B	A	C	B
Posttest	B	A	C	B	A	C
Delayed Posttest	C	B	A	C	B	A

Note. Letters refer to different video clips.

At each test, the researcher electronically distributed a prompt to each participant. Once 20 minutes had passed, the participants electronically sent their Microsoft Word file back to the researcher. The researcher circulated the classroom during the tests to ensure that the learners were following the instructions. These

tests were returned to the learners with comments addressing a range of language issues, but only after the delayed posttest had been completed.

Analysis

Test Score

The study measured the accuracy of the indefinite and definite articles together, following the majority of previous studies (e.g., Bitchener & Knoch, 2008; Sheen, 2007). In this regard, some studies did not include the definite article because its overgeneralized use may mask a learner's true L2 development (Shintani & Ellis, 2013; Shintani et al., 2014). While acknowledging their argument, the present study attempted to capture the overall development of the use of articles, expecting this measure to capture learners' ability to use articles in a comprehensible manner.

Accuracy for each target feature was measured by target-like use analysis (Pica, 1983). First, the obligatory occasions (i.e., contexts where the target feature must occur) for the definite and indefinite articles were counted. After that, the total number of correct provisions of indefinite and definite articles was tallied. Then, the instances of overgeneralized use of the two articles were also counted. The total number of correct provisions was divided by the sum of the obligatory occasions and the overgeneralization instances. Lastly, the values were multiplied by 100 in order for the test scores to be presented in the form of percentages. This method of scoring, by penalizing overgeneralized use, can minimize the potential inflation of the score that such use would otherwise bring. Furthermore, the scoring method has frequently been employed in past CF studies which investigated English articles (e.g., Sheen et al., 2009; Shintani & Ellis, 2013), and thus may help compare the results of the present study with the previous findings.

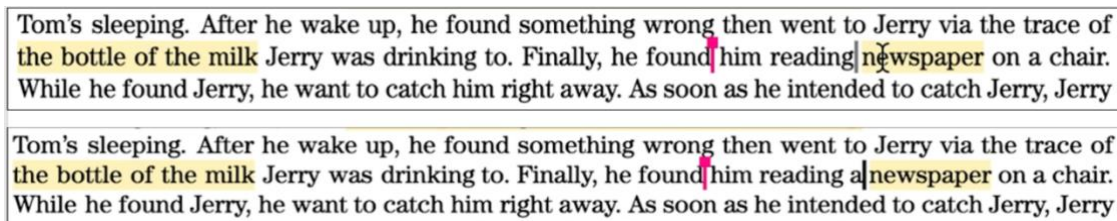
The researcher shared coding criteria with an L1 speaker of English, who was a doctoral student in Applied Linguistics, and they coded a few texts together as practice. Then, the two independently coded 18 texts (about 13% of the final data pool) that were randomly selected from the pretest and the posttests. The second rater was not informed of the source of each text (i.e., experimental vs. comparison, pretest vs. posttests). Intraclass correlation of the independent coding was .86, suggesting a good interrater reliability. Then, the raters discussed and solved all disagreements. Since some instances of the definite article turned out to be vague in terms of its function (e.g., uniqueness of a noun phrase), the raters decided to be conservative and to remove these cases to minimize potential inflation or deflation of scores. The researcher coded the rest of the data thereafter.

Revision Contribution

In order to track individuals' contributions to revision of a text, the screen capture data collected via QuickTime® during the treatment were manually examined. Since CF was provided only for the experimental group which consisted of 26 learners, only these 26 learners' computer activity was analyzed. Each screen capture data set consisted of approximately 40 minutes of a learner's on-screen operations. Given that one computer was available for each individual learner, and individual learners used their own institution accounts in Google Docs™, the screen capture data enabled the researcher to identify the agent of each particular revision. Specifically, in the screen capture data, a black cursor represented the learner who was using the computer, whereas a cursor in another color represented their partner or the researcher (Figure 3).

Figure 3

Revision Contribution Sample



Note. The black cursor is the learner who was using this computer, whereas the pink cursor represents his partner.

Revision contribution was operationalized as the production of the correct form in response to CF in Google Docs™ (Storch & Wigglesworth, 2010), regardless of the process the pair of learners undertook to produce the revision. Specifically, when a learner represented by the black cursor typed the correct form in response to CF, it was counted as one revision contribution for that individual learner, and no credit was given to that learner's partner.

Statistical Procedures

A series of statistical analyses was conducted on the test scores. For the first research question, the data were first checked for outliers. This screening process identified one potential outlier in the experimental group, a participant who used the definite article for all the non-pronominal noun phrases in the posttest, a pattern that deviated greatly from that of the other learners in the experimental group. Accordingly, this learner was removed from the data, and thus subsequent analyses were conducted for only 25 learners in the experimental group and 22 learners in the comparison group. First, normal distribution of the values for each of the six cells (the two groups' performance on the three tests) was confirmed by Shapiro-Wilk tests, and variance homogeneity was confirmed by Levene's tests. Mauchly's test of sphericity indicated that the assumption of sphericity was met. Then, a 3 (Time) \times 2 (Group) repeated measures Analysis of Variance (ANOVA) was performed. Time had three levels, namely pretest, posttest, and delayed posttest, whereas Group consisted of two levels, namely comparison group and experimental group. Effect sizes of post-hoc comparisons were calculated in the form of Cohen's *d*. The effect sizes were interpreted based on a benchmark proposed by Plonsky and Oswald (2014). Specifically, 0.60, 1.00, and 1.40 were considered as being small, medium, and large, respectively, for within-group differences, while 0.40, 0.70, and 1.00 were interpreted as being small, medium, and large, respectively, for between-group differences.

To explore the relationship between the number of revision contributions and learning, multiple regression analyses were performed. Each model had two independent variables, namely the pretest score and the number of revisions made for each individual. One model had the posttest score as a dependent variable, while the other model had the delayed posttest score as a dependent variable. Thus, two regression models were created. The pretest score was included as an independent variable to allow exploration of the relationship between revision contribution and posttest score, controlling for the effects of learners' baseline performance on their performance on the posttests. The regression for the posttest score had one data point whose absolute standardized residual was greater than 2.0 (i.e., -2.242), and the regression for the delayed posttest score included two data points whose absolute standardized residual was greater than 2.0 (i.e., 2.306, -2.208). However, these data points were included because their Cook's distance was minimal, being less than 1.0, suggesting that they had negligible influence on the prediction. The assumption of normality of residuals was reasonably met for each model based on the combined evidence from a Shapiro-Wilk test, a histogram, and a Q-Q plot. Durbin-Watson tests confirmed the independence of residuals, showing values from 1 to 3. The variance inflation factor indicated the absence of serious multicollinearity with values below 5. Finally, the collective evidence from Levene's tests and a scatterplot which plotted fitted values against standardized residuals indicated that variance homogeneity was reasonably met.

Results

CF Instances and Revision Contributions during the Treatment

Table 2 shows the number of CF instances for each pair of learners during the treatment and the number of revision contributions made by each learner in the pair. Table 2 illustrates considerable variation in the number of revisions produced by individual students.

Table 2

Instances of CF and Revision Contribution

Pair	CF	Individual	Revision Contributions
1	7	Student 1	3
		Student 2	3
2	4	Student 3	1
		Student 4	2
3	3	Student 5 ^a	1
		Student 6	0
4	9	Student 7	4
		Student 8	1
5	6	Student 9	0
		Student 10	1
6	2	Student 11	2
		Student 12 ^b	0
7	14	Student 13	2
		Student 14	4
8	7	Student 15	1
		Student 16	0
9	12	Student 17 ^a	9
		Student 18	2
10	17	Student 19	0
		Student 20	0
11	4	Student 21	0
		Student 22	2
12	10	Student 23	3
		Student 24	0
13	7	Student 25	3
		Student 26	0
14	5	Student 27	0
		Student 28	3

Notes. (a) Students who missed one of the test sessions; (b) Outlier.

For example, the students in Pair 1 contributed equally to the revision in terms of the number of their correct revisions in response to CF, whereas in Pair 9, Student 17 typed most of the revisions while Student 18 contributed much less. In contrast, neither student in Pair 10 responded to CF at all. These figures clearly demonstrate that providing CF to a pair of students does not guarantee evenly distributed contributions to revision from them. Furthermore, it is indicative that the provision of CF did not guarantee a response by learners, although the CF always included a short comment which prompted them to revise at a specific point in the text.

Effects of CF on Accuracy in a New Piece of Individual Writing

Table 3 sets out the descriptive statistics for the accuracy scores on the animation description tasks which each student worked on individually for the pretest, posttest, and delayed posttest. Overall, the experimental group's accuracy scores increased at each test, while the comparison group's accuracy scores decreased at the posttest and then increased at the delayed posttest.

Table 3

Descriptive Statistics for Accuracy Scores in the Tests

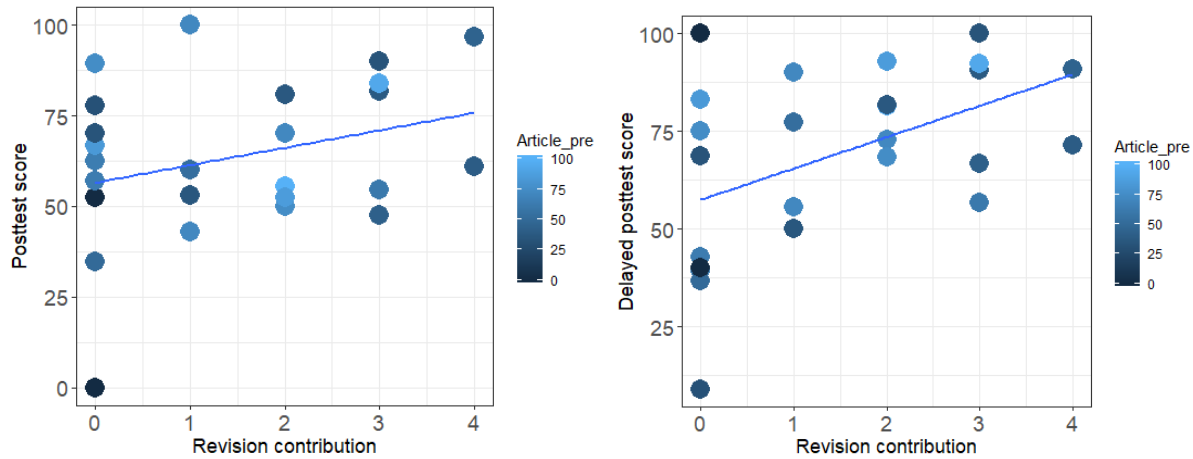
Group	n	M (SD)		
		Pretest	Posttest	Delayed Posttest
Experimental	25	54.08 (25.98)	63.62 (21.83)	69.34 (23.09)
Comparison	22	64.75 (25.72)	54.49 (21.87)	60.71 (18.46)

A 3 (Time) \times 2 (Group) repeated measures ANOVA was performed to identify statistically significant differences. No significant main effects were found for Time ($F(2, 90) = 1.142, p = .324, \eta^2_p = .025$), which suggests that the two groups taken together did not improve to a significant extent over the course of time. Also, the main effects turned out to be non-significant for Group ($F(1, 45) = .254, p = .617, \eta^2_p = .006$), indicating that the two groups did not differ notably. More importantly, however, the Time \times Group interaction effects were significant ($F(2, 90) = 3.293, p = .042, \eta^2_p = .068$). This means that there was a significant relationship between L2 development over the course of time and the presence of CF. The difference between the pretest and the posttest for the experimental group was less than small ($d = 0.40$), while the difference between their pretest and delayed posttest was barely small ($d = 0.62$). Similarly, the group difference was barely small for the posttest ($d = 0.42$) and the delayed posttest ($d = 0.41$), showing a similar pattern to that found in previous studies (Kang & Han, 2015).

Number of Revision Contributions As Predictor of Learning

The second research question concerned whether the extent of revision contributions predicts learning. Figure 4 shows scatterplots for the independent (i.e., revision contributions, pretest scores) and dependent variables (i.e., posttest scores, delayed posttest scores) in the two regression models. The x-axis shows the number of revision contributions, and the y-axis shows the score on the posttests. Pretest scores are represented by the depth of color. Figure 4 suggests there is a moderately positive relationship between the number of revision contributions and both posttest scores.

Figure 4
Scatterplot of Revision Contributions and Posttest Scores



The scatterplots seem to indicate that the more revisions individual learners made, the higher scores they obtained in both the posttest and the delayed posttest. Furthermore, given that data points of different depth of color are relatively evenly distributed, it seems that there is at most only a weak relationship between their pretest scores and posttest scores.

These overall preliminary patterns were partially confirmed by regression analyses. Specifically, the number of revision contributions did not predict the posttest scores to a significant extent ($p = .175$, $B = 4.40$ 95% CI [-2.11, 10.91]). Meanwhile, the number of revision contributions significantly predicted the delayed posttest scores ($p = .029$, $B = 7.74$, 95% CI [1.31, 14.18]). The adjusted R^2 is .18. This indicates that, for each correct revision a learner made, they scored higher in the delayed posttest by 7.74 points on average when pretest scores held constant, and that 18% of the variance in the delayed posttest scores is explained by the pretest scores and number of revision contributions made by individual students. The pretest score did not predict either the posttest score or the delayed posttest score. Table 4 summarizes the results from the regression analyses.

Table 4
Regression Analysis Summary

Dependent Variable	Predictor					
		<i>B</i>	95% CI for <i>B</i>	β	<i>p</i>	Adjusted R^2
Posttest	Pretest	0.18	[-0.17, 0.53]	.21	.300	.06
	Revision	4.40	[-2.11, 10.91]	.28	.175	
Delayed posttest	Pretest	0.11	[-0.23, 0.46]	.13	.501	.18
	Revision	7.74	[1.31, 14.18]	.47	.020	

Discussion

The first research question addressed the extent to which CF during computer-mediated collaborative writing is effective for improving grammatical accuracy in the use of the English articles in a new piece of individual writing. The results showed that the experimental group, which was initially outperformed by

the comparison group, improved in accuracy over the course of time and ended up outperforming the comparison group. In particular, their learning was salient in the difference between their pretest and delayed posttest scores (i.e., two weeks after the treatment). This result indicates the positive effects of CF in collaborative writing on grammatical accuracy in learners' subsequent independent performance. The fact that learners' performance was assessed through tasks based on different video prompts from the treatment video clips (although very similar) suggests that the accuracy improvement may not be merely a result of their chunk memorization but of their rule learning as well. Furthermore, this result appears to be in line with results from other written CF research, where indirect CF led to significant learning at an immediate posttest (Kang & Han, 2015).

Regarding the second research question, the present study hypothesized that the more revisions learners make, the more learning they demonstrate. Overall, the results seem to support this hypothesis given that a 1-unit increase in the number of revision contributions is related to approximately a 7-point increase in the delayed posttest score. This result appears to echo the findings in Shintani et al. (2014), where the difference between direct CF followed by revision opportunities and direct CF without revision opportunities emerged only at the delayed posttest. Although the present study withheld the correct form by adopting indirect CF, the explanation proposed by Shintani et al. (2014) may still be reasonable in this situation. That is, the present study may serve as further evidence that learners who revise their errors are likely to consolidate their declarative knowledge and thus facilitate their long-term learning (i.e., two weeks after). What distinguishes the present study from that of Shintani et al. (2014) is the comparable instructional time across the experimental and comparison groups. Accordingly, the present study may enhance the cogency of the case for the favorable effects of revision.

Interestingly, however, both the present study and Shintani et al. (2014) failed to find any impact of revision on participants' short-term learning. These findings possibly suggest that, even when learners do not have the opportunity to correct their errors, they may still notice the gap between their interlanguage and the target norm indicated by CF and achieve short-term learning. This speculation can be supported by written CF studies which have reported positive effects of CF without having learners correct their errors (e.g., Sheen, 2007; Stefanou & Révész, 2015). Another possibility is that learners benefited from observing their peers' revision. In the interface of Google Docs™, learners are able to observe what their partners are writing, including their responses to instructors' CF, in a synchronous manner. In fact, previous studies found that learners appreciated the revision history affordance available in a computer-mediated collaborative writing platform and benefited from observing how other group members dealt with their instructor's feedback (Lee, 2010). The other possibility is that, since this study paired students, they were afforded opportunities to consult with their peers about CF. This peer interaction may have functioned as an opportunity for those who did not have a chance to make corrections by themselves to achieve short-term learning. However, this pair interaction is beyond the scope of the present study.

Conclusion

The present study investigated the effects of CF in in-class computer-mediated collaborative writing on the accurate use of English definite and indefinite articles in a new piece of individual writing. Overall, the present study offered additional evidence of the positive effects of CF in collaborative writing, suggesting the transferability of learning in collaborative writing to an individual's L2 development, manifested in a new piece of writing. Also interesting is that revision contributions may be related to long-term learning. These findings are important in several ways in both teaching and research. First, instructors may be encouraged to design tasks in such a way that all group members need to contribute to revision activity so that each member can experience proceduralization of their linguistic resources. In the implementation, instructors may need to make moment-by-moment efforts to have individual learners contribute to the revision, by, for example, having one group member take the lead in revising their text if another dominates the revision activity. Computer-mediated platforms enable instructors to track the contributions of each

group member (Yim & Warschauer, 2017). For example, emerging collaborative writing platforms, such as Quip and Etherpad, have great potential in this regard. This suggestion further pertains to methodological issues in previous research, which often reported revisions per pair without tracking the individual agent of a particular revision. In short, studies may be encouraged to report revision per individual in order to avoid overestimating the learning of individuals who rarely contribute to a group revision.

However, these arguments and recommendations are tentative, due to some limitations of the study. First, the convenience sampling limited the ability to control biographical variables (e.g., age, L1). In particular, it should be noted that the participants were fully matriculated at a U.S. college with plentiful opportunities to be exposed to English on a daily basis. Therefore, it is questionable to what extent the current findings are generalizable to other proficiency levels and contexts where learners are rarely exposed to their target language. Second, the present study employed a so-called one-shot treatment and this may have limited the opportunity to observe a clearer L2 development which learners might have achieved with a few more treatment sessions. This one-shot design may also have minimized transfer of learning during the treatment to each individual if we assume that learners were not familiar with the given task and their partner enough to engage in collaborative dialogue. In fact, a recent study found that learners gradually become able to create a collaborative atmosphere through task repetition (Chen, 2019). On the other hand, however, this one-shot design may have helped to illustrate a clear association between revision contributions and learning while more treatment sessions might have had learners achieve high accuracy scores to the extent which does not manifest a score variation (Révész et al., 2011). Third, the present study did not analyze peer interactions during the treatment. Such analysis could have shed light on learners' overt collaboration on making revisions orally, which is not well assessed by counting revisions produced in Google Docs™. Fourth, the single type of test score limited the representativeness of the learners' performance, leaving unanswered any questions as to whether learners developed different types of knowledge. These limitations render the arguments based on the present study tentative, and future research is called for before a solid conclusion can be drawn.

Despite these limitations, the present study illustrated the favorable effects of CF in collaborative writing on individuals' L2 development. Future research can expand the present study in several ways. For example, it would be interesting to examine whether the effects of CF vary depending on the participatory structure (e.g., individual, pair, group). In particular, it would be useful to examine the effects of CF in group work, aiming for a time-saving implementation of written CF for a large class. Furthermore, future research, perhaps adopting a laboratory setting, could examine whether learners benefit from observing their fellow group members' revision. In fact, recent studies have empirically pursued this kind of inquiry for CF given orally (Rassaei, 2019; Yilmaz, 2016). If this was empirically tested for written CF, it would provide evidence regarding the validity of the claim that learners who make fewer revision contributions learn less. Future studies could also examine the generalizability of the findings to other writing tasks which involve a higher degree of knowledge transformation (e.g., writing argumentative texts) (Scardamalia & Bereiter, 1987), given that these tasks may not draw learners' attention to grammar as much as descriptive writing prompts. Equally importantly, given the reported relationship between revision contributions and learning, researchers conducting future studies in written CF and collaborative writing are strongly encouraged to report not only on whether the study asked learners to revise their errors but also on to the extent to which learners *actually* revised the errors and how revision work was distributed in a pair/group. This kind of data report may advance the field by identifying task process factors responsible for varying L2 development as a result of the same task workplans.

Acknowledgments

This study was conducted as part of the portfolio assessment for the PhD program in Applied Linguistics and Technology at Iowa State University. I am sincerely grateful to the committee members for their constructive feedback on the previous version of this manuscript. I also greatly appreciate the anonymous

reviewers' constructive feedback. I am also appreciative of the students' participation in this study and Dr. Jordan Smith's willingness to be a second rater. Last but not least, I would like to extend my utmost gratitude to Dr. Jim Ranalli for his advice at various stages of this study. All errors are my own.

References

- Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, 7(3), 347–376.
- Ammar, A., & Hassan, R. M. (2018). Talking it through: Collaborative dialogue and second language learning. *Language Learning*, 68(1), 46–82.
- Benson, S., & DeKeyser, R. (2019). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, 23(6), 702–726.
- Bikowski, D., & Vithanage, R. (2016). Effects of web-based collaborative writing on individual L2 writing development. *Language Learning & Technology*, 20(1), 79–99. <http://dx.doi.org/10125/44447>
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102–118.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12(3), 409–431.
- Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles: An analysis of the metalinguistic knowledge used by Japanese students in acquiring the English article system. *Studies in Second Language Acquisition*, 24(3), 451–480.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267–296.
- Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, 12(1), 43–64.
- Chen, W. (2019). Patterns of pair interaction in communicative tasks: The transition process and effect on L2 teaching and learning. *ELT Journal*, 72(4), 425–434.
- Coyle, Y., & de Larios, J. R. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition*, 36(3), 451–485.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge University Press.
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 94–112). Routledge.
- Ellis, R. (1990). *Instructed second language acquisition*. Basil Blackwell.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161–184.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Korama.
- 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), 1–18.

- Karim, K., & Nassaji, H. (2020). The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Language Teaching Research*, 24(4), 519–539.
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183–199.
- Kuiken, F., & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, 37(3–4), 343–358.
- Lee, L. (2010). Exploring wiki-mediated collaborative writing: A case study in an elementary Spanish course. *CALICO Journal*, 27(2), 260–276.
- Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Multilingual Matters.
- Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, 30, 66–81.
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, 25(2), 215–232.
- McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32(2), 207–224.
- Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning*, 33(4), 465–497.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912.
- Rassaei, E. (2019). Effects of dynamic and non-dynamic corrective feedback on EFL writing accuracy during dyadic and small group interactions, *International Review of Applied Linguistics in Language Teaching*. Advance online publication. <https://doi.org/10.1515/iral-2016-0044>
- Reinders, H. (2009). Learner uptake and acquisition in three grammar-oriented production activities. *Language Teaching Research*, 13(2), 201–222.
- Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and L2 development. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 203–236). John Benjamins.
- Riazi, M., Shi, L., & Haggerty, J. (2018). Analysis of the empirical research in the journal of second language writing at its 25th year (1992–2016). *Journal of Second Language Writing*, 41, 41–54.
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, Vol. 2. Reading, writing, and language learning* (pp. 142–175). Cambridge University Press.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556–569.

- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20*(4), 286–305.
- 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*(1), 103–131.
- Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *The Modern Language Journal, 99*(2), 263–282.
- Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System, 27*(3), 363–374.
- Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research, 5*(1), 29–53.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*(1), 119–158.
- Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research, 11*(2), 143–159.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.
- Storch, N., & Wigglesworth, G. (2007). Writing tasks: Comparing individual and collaborative writing. In G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Multilingual Matters.
- Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition, 32*(2), 303–334.
- Swain, M. (2006). Language, agency and collaboration in advanced language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). Continuum.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research, 37*(3-4), 285–304.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327–369.
- Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing, 8*(2), 111–122.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255–272.
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning, 62*(1), 1–41.
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26*(3), 445–466.
- Wigglesworth, G., & Storch, N. (2012a). Feedback and writing development through collaboration: A socio-cultural approach. In R. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 69–101). De Gruyter Mouton.
- Wigglesworth, G., & Storch, N. (2012b). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing, 21*(4), 364–374.

- Yilmaz, Y. (2016). The role of exposure condition in the effectiveness of explicit correction. *Studies in Second Language Acquisition*, 38(1), 65–96.
- Yim, S., & Warschauer, M. (2017). Web-based collaborative writing in L2 contexts: Methodological insights from text mining. *Language Learning & Technology*, 21(1), 146–165.
<https://dx.doi.org/10125/44599>
- Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research*, 24(2), 227–250.
- Zdorenko, T., & Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, 32(1–2), 38–62.
- Zhang, M. (2019). Towards a quantitative model of understanding the dynamics of collaboration in collaborative writing. *Journal of Second Language Writing*, 45(3), 16–30.

About the Author

Taichi Yamashita holds a PhD in Applied Linguistics and Technology from Iowa State University. His research interests address the intersection of CALL and instructed SLA with specializations in feedback on writing and collaborative learning. His articles have appeared in *Foreign Language Annals*, *Computer Assisted Language Learning*, and *Applied Linguistics*.

E-mail: taichiy@iastate.edu