

The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article

Natsuko Shintani and Rod Ellis

University of Auckland

Abstract

The study extends current work on written error feedback in writing in two ways. First, it examines whether it has an effect on adult ESL learners' L2 implicit and explicit knowledge. Second, the study compares the effect of one common type of feedback— direct corrective feedback (DCF) - with an alternative type of error feedback - the provision of metalinguistic explanation (ME). The effect of these two types of error feedback was measured by an Error Correction Test (ECT) and by examining the accuracy of use of the target feature (the English indefinite article) in both a revised text and in new pieces of writing by 39 low-intermediate ESL students in an intensive language programme in the United States.. In addition, eye-tracking data and self-reports elicited from the learners provided information about the use that they made of the DCF and ME. It was found that the DCF had no effect on accurate use of the target feature suggesting that it benefitted neither implicit nor explicit knowledge. In contrast, the ME led to gains in accuracy in the ECT and in a new piece of writing completed immediately after the treatment but not in a second new text completed two weeks later. These results are interpreted as indicating that the ME helped to develop learners' L2 explicit knowledge but that the effect was not durable and thus probably had no effect on their implicit knowledge. Learners' self-reports indicate that the learners receiving the DCF did not develop awareness of the rule whereas those receiving the ME did and were able to use it when revising their original text. These findings are discussed from the perspective of both SLA theory and language pedagogy and suggestions for further research are put forward.

Error correction and implicit/ explicit L2 knowledge

A recent special issue of the *Journal of Second Language Writing* (Volume 21, 4) was devoted to exploring the interface between L2 writing research and second language acquisition (SLA). Contributors to this special issue pointed out that from an SLA perspective a key question is whether written error feedback contributes only to learners' explicit L2 knowledge or also to their implicit L2 knowledge. Acknowledging that the existence of these two types of knowledge is generally accepted in SLA, Williams (2012) proposed that the crucial issue was

whether error feedback simply encourages learners to tap into their explicit knowledge or whether it also facilitates “L2 development” (i.e., their implicit knowledge). Bitchener (2012) was similarly concerned with whether error feedback had any effect on implicit knowledge. Polio (2012) argued that “establishing a research agenda on the roles of explicit and implicit knowledge in writing is crucial” (p. 408). The study reported below is intended to make a start on such an agenda.

The implicit/ explicit issue lies at the heart of the debate about the efficacy of error correction, and, also perhaps can help explain a misunderstanding about what this debate is actually about. The debate has been ongoing since Truscott’s (1996) article in *Language Learning*, in which he claimed there was neither any empirical or theoretical justification for correcting students’ written errors. In a series of further articles (Truscott 1999, 2004, 2007, 2010) he continued to reject any role for error correction where grammar is concerned (although not for other aspects of writing). The nub of Truscott’s argument is that error correction will have no effect on the development of the type of knowledge needed to engage in writing or speaking for communicative purposes because it is impossible to know which grammatical features learners are developmentally ready to acquire. However, he acknowledged that it could have an impact on the explicit knowledge needed for monitoring in grammar tests or when revising a written text that has been corrected. Truscott’s perspective, then, is that of an SLA theorist. In contrast, the defense of error correction mounted by Ferris (1999, 2006) drew not on SLA but on L2 writing pedagogy. The nub of her argument, since supported by several empirical studies (see Bitchener & Ferris, 2011, for a review of these) was that error correction is a valuable pedagogic tool because it can help learners to write more accurately. These positions appear oppositional but they need not be seen in this way. Truscott may be right in asserting that error correction does not contribute to what he called “genuine knowledge of language” (Truscott, 1998, p. 120) – i.e., implicit knowledge. However, the empirical evidence indicates that he is wrong in claiming that it has no effect on accuracy in writing for communicative purposes. As Williams (2012) noted, two “inherent features of writing” (its permanence and the fact that it does not have to be produced online) “permit the learner more control over attentional resources as well as more need to attend to language both during and after production” (p. 322). In other words, error correction may be effective because it assists learners’ explicit knowledge which they can then use to eliminate errors in their writing.

Interestingly, the contributors to the special issue of the *Journal of Second Language Writing* were also of the view that error feedback contributes to explicit rather than implicit knowledge. Polio (2012), for example, commented “I have assumed ... that written corrective feedback will increase explicit knowledge” (p. 386). Bitchener (2012) argued that the result of the studies that have investigated written error correction “provide evidence that it is possible for learners to internalize and consolidate their explicit knowledge” (p. 353) and later “we do know

that written CF can play a role at least in terms of developing explicit knowledge” (p. 361). What is, therefore, an “assumption” for Polio is “knowledge” for Bitchener. In fact, all that we do know is that written error correction leads to improved accuracy in writing. We do not know whether this is because it benefits explicit knowledge or “genuine knowledge of language” for, as Polio pointed out, learners can use both types of knowledge when writing and there have been no studies that have specifically addressed the effects of written error correction on explicit knowledge or, we would add, on implicit knowledge.

Given the lack of empirical evidence about the effects of written error feedback on explicit and implicit knowledge, it is necessary to turn to SLA theories and to research on the effects of oral corrective feedback. Here we find a number of different positions, which can be summarized as follows:

1. Written corrective feedback is inevitably explicit in nature and, like other forms of explicit instruction, will only have an effect on learners’ explicit knowledge. This is the position that Truscott has taken, along with Krashen (1982) and Schwartz (1993). It assumes that there is no interface between explicit and implicit knowledge.
2. Oral feedback that provides learners with positive evidence (for example, recasts) can contribute directly to the development of implicit knowledge (Long, 1996). Direct written corrective feedback is analogous to recasts in this respect but also differs from it in an important way. Oral recasts occur in online interaction when meaning or form is negotiated whereas direct written corrective feedback occurs offline (i.e., there is typically a delay between the time the error is made and the feedback received). Cognitive-interactionist theories consider that for feedback to be effective it needs to be provided online in a “window of opportunity” (Doughty, 2001). Thus, even written feedback that affords positive evidence may not impact on learners’ implicit knowledge.
3. Corrective feedback that results in explicit knowledge may indirectly contribute to the development of implicit knowledge by promoting “noticing” and “noticing-the-gap” (Schmidt, 1993). This position is compatible with the weak interface position (R. Ellis, 1993; N. Ellis, 2005). However, the interface between explicit and implicit knowledge does not occur immediately as implicit knowledge is only developed gradually as learners are exposed to more L2 input and opportunities for production. This would suggest that one-shot written error feedback of the type investigated in many studies (including the one we report below) may contribute to explicit knowledge but will have no effect on their implicit knowledge.
4. Corrective feedback assists skill-learning (DeKeyser, 1998) if it “pushes” learners to self-correct. From this perspective, output-prompting feedback that results in learners correcting their own errors will be more effective than input-providing feedback as Lyster (2004) found. Applied to written error correction, this theory points to the importance of

learners' having the opportunity to revise their writing following feedback. Correcting errors in the process of revising can be viewed as a type of contextualized practice that helps learners proceduralize their explicit knowledge.

There is now ample empirical evidence (see Ellis & Shintani, 2013, for a review of the relevant studies) that oral corrective feedback leads to improved accuracy in oral free-production tasks, which for many SLA researchers provide the best evidence that "real acquisition" (as opposed to "pseudo-acquisition") has taken place. However, given the differences in the nature of oral and written corrective feedback, it cannot be assumed that the latter will have the same effect as the former. Nevertheless, as shown above, there are theoretical grounds for arguing that it might do so.

To sum up, while there is clear support for the claim that written error feedback can help learners to write more accurately, there is no study to date that has attempted to investigate whether it assists the development of explicit knowledge, the acquisition of implicit knowledge or both. There is, however, an assumption that it will assist explicit knowledge. There are also theoretical reasons for believing that either indirectly or directly it can contribute to the acquisition of implicit knowledge although, as Truscott has consistently argued, there are also theoretical grounds for disputing this. As Ortega (2012) pointed out, the test of these different theoretical positions faces a number of "methodological hurdles" to do with how to determine which of two types of knowledge a learner is using. We will return to this issue later. First, though, we will introduce the two types of written error correction we investigated and our reasons for choosing them.

Direct written corrective feedback and metalinguistic explanation

Direct corrective feedback (DCF) and metalinguistic explanation (ME) constitute two different ways of carrying out written error correction. DCF provides learners with the correct forms but it is a time-consuming endeavor as each learner's written work needs to be corrected and it requires little depth of processing by learners. In contrast, giving all the students the same ME in the form of a handout following writing is much less time-consuming as the teacher does not need to correct each learner's written work. Also, arguably, it encourages greater depth of processing on the part of the students as they have to apply the ME to their own errors. However, it assumes that students are able to use the ME to identify and self-correct their own errors. To date, despite the large body of studies that have investigated CF (see Bitchener & Ferris, 2011), no study has compared these two approaches for addressing errors in learners' writing ^[1].

Direct written corrective feedback

A broad distinction can be drawn between indirect and direct CF. In indirect CF errors are indicated but no corrections are provided. We have elected to investigate direct CF as we wished

to maximize the potential effect of the CF. Direct CF is more likely to facilitate learning when learners have no or ill-formed explicit or implicit knowledge of a grammatical feature. Chandler (2003) argued that direct CF enables learners to internalize the correct form immediately whereas indirect CF does not inform learners about the accuracy of their hypothesized corrections. This is especially the case if learners have limited L2 proficiency and poorly developed writing abilities as was the case in the learners we investigated.

A number of studies have investigated the effects of direct CF. Ferris and Robert (2001) and Chandler (2003) reported that direct CF was effective in enabling learners to revise their errors. However, this is not surprising as they had access to the corrections when revising so they only needed to copy the corrections. More important, as Truscott has repeatedly pointed out, are studies that have investigated the effect of CF on grammatical accuracy in new pieces of writing as these can afford evidence of whether any learning has taken place. Early research comparing direct and indirect CF produced conflicting results (see Bitchener & Knoch, 2008, for a review). Later studies pointed to an advantage for direct CF. Bitchener and Knoch (2010), for example, found both types were equally effective over the short term but only direct CF had a sustained effect. Van Beunigen *et al.* (2012), in a carefully designed study that took account of differences in learners' initial proficiency, reported that unfocused direct and indirect CF were equally effective for non-grammatical errors but that only direct CF resulted in significant gains in grammatical accuracy. Other studies (e.g., Ellis *et al.*, 2008) which investigated only direct CF have shown that it improves accuracy in both the short and the long term. Farrokhi and Sattapour (2012) also reported that direct corrective feedback resulted in significant gains in accuracy in the use of English articles in a new piece of writing. Only one recent study (Truscott & Hsu, 2008) failed to show that direct CF improved accuracy in new writing^[2]. Overall, the research indicates that direct CF is effective not only in assisting revision but also in improving accuracy in new pieces of writing. Furthermore, it shows that the effects of the CF are evident not just immediately following the CF treatment but in the long-term. This suggests that it has contributed to change in the learners' implicit L2 knowledge.

We also elected to investigate focused CF, where only one type of error is corrected and the rest ignored, rather than unfocused CF, where all (or most) errors are corrected. This decision was motivated by recent studies (e.g., Bitchener & Knoch, 2008, 2009, 2010) which have shown that focused CF is effective and by Farrokhi and Sattapour's (2012) study, which showed that it was more effective than unfocused CF^[3]. It was also motivated by the fact that in order to compare CF with ME it is necessary to pre-select a specific grammatical feature for attention. However, we acknowledge that there are arguments in favor of unfocused CF. Van Beunigen *et al.* (2012) suggested that the writing tasks in focused CF studies may have resembled grammar exercises rather than authentic writing tasks and thus may have led learners to consciously monitor use of the target feature^[4]. Van Beunigen *et al.* (2012) also argued that unfocused CF

accords more with normal pedagogic practice as “a teacher’s purpose in correcting his/her pupils written work is to improve accuracy in general, not just the use of one grammatical feature” (p. 6). Despite these arguments, we consider focused CF to be pedagogically sound. Teachers can vary the feature they focus on in different writing tasks and thus achieve a wide coverage of grammar over time. Also, focusing on a single grammatical feature enhances the likelihood that learners will not just memorize the specific corrections but develop an awareness of the underlying rule (i.e., develop explicit knowledge). This is much less likely to occur if many different kinds of errors are corrected. Andersen (2010), for example, investigated “tiered feedback” (i.e., feedback that began by focusing on one grammatical feature and then proceeded to add an additional feature each time feedback was provided) and found that the effectiveness of the feedback lessened as the number of corrected features increased. There is also a methodological advantage in examining focused CF. As Van Beunigen *et al.* acknowledged, in unfocused CF the specific structures corrected in an initial piece of writing may not occur in subsequent writing. To demonstrate that CF has an effect on either explicit or implicit knowledge it is necessary to show that the correction of a specific error leads to the elimination of or reduction in occurrences of that error. This is much easier to achieve in focused CF.

Metalinguistic explanation

The type of metalinguistic explanation we investigated differs from the type investigated in previous error correction studies. In previous studies, ME has been provided by means of error codes indicating the type of error a learner has made (as in Robb, Ross & Shortread, 1986) or by numbering errors and then providing a brief metalinguistic explanation of each type of error (as in Sheen, 2007). In both cases, it is necessary to identify and correct the errors in individual learners’ written work. The ME we investigated took the form of a handout providing an explanation of the target structure (articles), which was given to all the students when they had finished writing. Thus, no correction of individual learners’ writing took place. To the best of our knowledge no previous study has investigated this type of ME.

The ME in our study can be seen as a form of direct consciousness-raising (Ellis, 1997). A number of studies have shown that this is effective in developing learners’ explicit knowledge. Fotos and Ellis (1991) compared ME with indirect consciousness-raising involving a grammar discovery task. They reported that both types of consciousness-raising resulted in significant gains in learners’ understanding of the target structure (dative alternation) as measured by a grammaticality judgment test. However, the ME led to more durable gains. In a follow-up study that investigated four different grammatical structures, Fotos (1994) also found ME to be effective. In both of these studies the explicit metalinguistic explanation was provided orally by a teacher. However, Mohamed (2001) used handouts in her comparative study of ME and indirect consciousness-raising. She found that for low-intermediate learners similar to the learners we

investigated ME was more effective. However, the ME in our study differed from that in these studies as it was provided as feedback on learners' written work rather than as pre-emptive, stand-alone explanation.

Investigating learners' response to instruction

The majority of the studies cited above have investigated the effect of CF or ME in terms of learning outcomes (i.e., gains in accuracy). They did not examine the learners' response to the feedback they received or their use of it when rewriting. We argue that if we are to investigate whether written error feedback leads to explicit knowledge we need to find out how learners respond to the feedback. In the case of DCF we need to know whether learners attend to the corrections provided and also whether they use the corrections to try to understand the rule that has been broken. In the case of ME we need to find out if learners have understood the explanation and whether they were able to apply it by identifying and self-correcting the errors they have made. What is needed, then, are methods for investigating how learners have responded to the correction and the use they have made of it.

There have been very few studies of learners' actual engagement with written correction. Sachs and Polio (2007) used think-aloud and interviews to examine how learners responded to error correction. They found that "reports of noticing during the processing of written corrective feedback were related to subsequent revision changes" (p. 85). Researchers investigating oral corrective feedback have paid more attention to how learners process feedback by using stimulated recall to elicit learners' retrospective comments on specific corrections (e.g., Egi, 2007). In our study we used eye-tracking to investigate where and for how long learners fixated on their article errors. We also carried out interviews and stimulated recall to investigate the extent to which the learners developed explicit knowledge of the target feature and what use they made of it when revising.

Measuring the effects of written error correction on learners' implicit and explicit knowledge

There are no "pure" measures of L1 implicit and explicit knowledge (Ellis, 2009). This is because, under any conditions of use, learners will make use of the linguistic resources at their disposal. However, as Ellis pointed out, it is possible to design instruments that will bias learners to the use of one type of knowledge or the other. He identified four criteria that could be used to design instruments for measuring the two types of knowledge:

1. Degree of awareness (i.e., whether the instrument favors the use of "feel" or "rule")
2. Time available (i.e., whether the instrument puts pressure for learners to process online or

Commented [JSLW1]: Please unpack.

Commented [JSLW2]: You may want to add Wigglesworth & Storch 2010, in SSLA, to the research mentioned here.

whether it allows for offline processing)

3. Focus of attention (i.e., whether the instrument focuses learners' primary attention on meaning or on form).
4. Utility of knowledge of metalanguage (i.e., whether the instrument requires or induces learners to access their knowledge of metalanguage).

In the study reported below two instruments were used – an Error Correction Test (ECT) and Narrative Writing Tasks.

The ECT was chosen to afford a measure of learners' explicit knowledge. This test, which required learners to identify errors in sentences and then write out the sentences correctly, clearly favors the use of "rule" (although learners could also respond by "feel"), was unpressured, required a primary focus on form, and potentially made the use of metalanguage advantageous. Tests such as the ECT have not been typically used in written error correction studies where the main purpose is to investigate whether correction has any effect on accuracy in new pieces of writing (but see Ellis *et al.*, 2008). Clearly, such a test cannot be used to address whether error correction improves writing. As Polio (2012) pointed out most writing researchers "would argue that the improvement has to be on a piece of writing as opposed to a grammar test" (p. 377). However, such a test can provide evidence of whether the correction helped to develop learners' explicit knowledge and that was its purpose in our study.

The Narrative Writing Tasks consisted of picture compositions; learners' were instructed to write out the story depicted in the pictures and were given a limited amount of time to do so. Such a task favors the use of "feel" (although learners can also try to apply "rules"), was time-pressured, required a primary focus on meaning and did not encourage learners to access metalanguage. Thus, it can be considered as biasing learners to the use of their implicit knowledge. However, as Polio (2012) noted "any writing task, even those completed under time constraints, can tap into both explicit and implicit knowledge (p. 377)". Thus, the validity of the Narrative Writing Task as a measure of implicit knowledge is doubtful. However, its validity can be substantiated in two ways: (1) by including a delayed writing task in the design of a study and (2) by obtaining evidence of how learners oriented to the task. Implicit knowledge is durable (i.e., once acquired it is not easily lost) whereas explicit L2 knowledge, like other types of declarative knowledge, can be easily forgotten. Thus, demonstrating that any improvement in accuracy is durable provides support for claiming that correction has an effect on learners' implicit knowledge. Eliciting self-reports from learners to ascertain whether they saw the writing task as an opportunity to practice grammar or free writing and whether they paid conscious attention to the target form when writing can also help to confirm whether or not there was a primary focus on meaning. In the study we now report, we included a delayed writing task and also elicited self-reports from the learners by means of an Exit Questionnaire.

Research questions

The study had two main purposes. It sought to distinguish the effect of form-focused feedback on the acquisition of learners' implicit and explicit L2 knowledge. It also sought to compare the effect of a type of form-focused feedback (metalinguistic explanation) not previously investigated with that of a well-researched type (direct corrective feedback).

RQ1: What effect does focused direct corrective feedback (DCF) have on learners' acquisition of explicit and implicit knowledge of the English indefinite article?

RQ2: What effect does metalinguistic explanation (ME) have on learners' acquisition of explicit and implicit knowledge of the indefinite article?

RQ3: Is there any difference in the effect that direct focused corrective feedback (DCF) and metalinguistic explanation have on the acquisition of implicit and explicit knowledge of the indefinite article?

To answer these research questions L2 explicit knowledge of the target feature (indefinite article) was measured by means of an Error Correction Test. L2 implicit knowledge was measured by examining the use of the target feature in new pieces of writing.

RQ4: What use did individual learners make of a) the DCF and b) the ME when revising their text?

To address this research question we examined the frequency and duration of six learners' eye fixations when they studied their original piece of writing after receiving feedback. We also conducted an oral interview with the same six. The purpose was to investigate how they handled specific corrections and to determine the extent to which they had become aware of the rule for the use of the indefinite article.

Method

Participants

The study involved five intact low-intermediate academic writing classes in an Intensive English Language Program in the United States. A total of 49 students who signed the consent form approved by the researchers' university participated in the study. There were 31 male and 18 female students aged 18 to 48 (22.6 years old on average). The students had been living in the US for between three and 18 months but had also studied English as a foreign language in their own countries. Their first languages were Arabic (n=34), Chinese (n=14) and Korean (n=1). It

Commented [JSLW3]: Why? Please explain the connection with the University of Auckland.

was clear from the pre-test writing task that the participants' writing ability was limited both in terms of their linguistic knowledge and their familiarity with English orthographic conventions.

The participants were taught in five intact classes, which were randomly assigned as the direct corrective feedback (DCF) group (class 1, n=15), the metalinguistic explanation (ME) group (class 2, n= 13), the control group (class 3, n=6 and class 4, n=9). The students in Classes 1, 2, 3, and 5 were classified by the institution as "Level 1" and Class 4 (a part of the control group) as "Level 2". Six additional students from class 5 were recruited for eye-tracking and an interview. These students, therefore, came from the same class level as the students in the main feedback study.

Target structure – indefinite article "a/an"

The target structure was the use of the English indefinite article to denote a specific referent unknown to the hearer (i.e., + Specific Referent/ - Assumed to Hearer) as in "There was **a** tall girl standing in front of me".

As Young (1996) noted:

The problems second-language learners face in mastering the English article system are indeed tough. They include sorting out many-to-many form-function mappings, learning the properties of nouns and how they change with context, and determining ways in which to mark specific/ nonspecific reference and shared/ unshared context (p. 142).

The essential problem lies in the fact that the multi-functional nature of articles does not accord with the One to One Principle (Andersen, 1984), which guides learners in early L2 acquisition.

The choice of the indefinite article was motivated by studies which have shown that it is frequently late acquired. Huebner (1983) reported that the learner he investigated began by using "the" for all functions and only started to use "a" later. Young's (1996) study of Czech and Slovak learners of English showed that the low proficiency learners (of a similar type to those we investigated) only marked 7% of the contexts requiring "a/an" for first mention correctly (the high proficiency learners only managed 20%).

A number of recent CF studies have investigated articles (e.g., Bitchener & Knoch, 2008, 2009). They examined both "a" and "the" for first/ anaphoric mention. However, learners' tendency to overgeneralize the use of the definite article makes it difficult to determine if they have really acquired "the" to realize a specific grammatical function. Restricting the analysis to "a" for first mention allows for a more reliable scoring of the effect of instruction on acquisition.

Design

Three groups (the DCF, the ME and the control) participated in three sessions: Time 1, Time 2 (two days after Time 1) and Time 3 (two weeks after Time 2). At Time 1, they completed a background questionnaire, the Error Correction Test, and the first writing task. At Time 2 the experimental groups received their respective feedback and were then asked to revise their original writing followed immediately by a new piece of writing. At Time 3, two weeks later ^[5], the groups completed a third piece of writing, the exit questionnaire and the same Error Correction Test as at Time 1. The researcher informed the learners' normal teachers about the focus of the study but asked them not to teach articles during this research project.

In order to eliminate any possible task effect, the three writing tasks were counterbalanced; each group was divided into three sub-groups and each sub-group completed a different task at times 1, 2 and 3 as shown in Figure 1 below.

[INSERT FIGURE 1 HERE]

Instruments

Writing task materials. Three picture composition tasks from Heaton (1975) were used to elicit stories from the students. Each picture sheet consisted of six pictures that describe a short story (see Appendix A for an example). The participants were given 20 minutes to complete the task which, given their limited proficiency, posed considerable pressure on their language processing capacity. They were allowed to use their electronic dictionaries to assist them with vocabulary as it was anticipated that they would not know some of the vocabulary needed to write the stories. Otherwise they completed the stories without assistance.

The writing tasks were designed to afford a potential measure of learners' implicit knowledge of the indefinite article. In accordance with Ellis' (2005) criteria for tests of implicit knowledge (see above), the writing tasks did not require awareness of the target form, the tasks were pressured, the primary focus was on meaning (i.e., telling the story) and the task instructions did not invite the use of metalinguistic knowledge. However, the pre-writing task more obviously satisfied these criteria than the post-writing tasks which were completed after the learners had received the feedback. The feedback can be seen as inviting conscious attention to the indefinite article. Also, as already noted writing by its nature allows greater opportunity for conscious monitoring of linguistic forms. Nevertheless, the primary focus remained on meaning in the post-writing tasks and it was clear that many of the learners continued to struggle to complete the tasks in the 20 minutes allocated. Thus, while these tasks made use of explicit knowledge of the target feature possible, opportunity for the kind of controlled processing this requires was possible but clearly limited.

Error Correction Test. The test (see Appendix B) was the same test as in Ellis *et al.* (2008). It consisted of 16 decontextualised sentences each containing a single error. Six sentences

included errors in the use of the indefinite article for first mention and the other 10 included other grammatical errors (e.g., definite article, verb tense, and 3rd person *-s*). The test was completed before the participants started their first piece of writing and again at the end of the study after the final piece of writing. In this way it served as a pre-test and post-test. The reliability of the test was assessed by Cronbach's alpha. The value obtained was 0.79, which can be considered acceptable for a test with relatively few items.

The purpose of the test was to provide a measure of learners' explicit knowledge of the indefinite article. The test required learners to (1) identify the error in each sentence and (2) write out the sentence correcting the error. This test satisfied all the four criteria for a test of explicit knowledge proposed by Ellis (2009). That is, it required a high degree of awareness on the part of the learners, there was no time pressure, it focused attention on form rather than on meaning and it potentially drew on learners' metalinguistic knowledge of the indefinite article rule. In effect, the test was a measure of learners' ability to monitor, which according to Kormos (2000) involves conscious decision-making.

Background questionnaire. The purpose of this was to provide a profile of the individual students. It requested information about the participants' age, gender, mother tongue, the length of time they had spent in the US, and the length of time they had studied English in their home country.

Metalinguistic explanation handout. The handout consisted of an explicit explanation of the use of definite and indefinite articles. The complete handout is in Appendix C.

Exit questionnaire. This questionnaire involved one multiple choice question and an open-ended question that asked the participants' perceptions about the goal of this study and their focus while they were writing the tasks. All the groups completed this questionnaire at the end of Time 3.

Procedures

Direct corrective feedback group. The direct corrective feedback (DCF) involved indicating the location of any article error on the student's text and then providing the correct form or inserting a missing article. Both definite and indefinite article errors were corrected. The DCF group received direct corrective feedback only on the writing they produced at Time 1 (see Appendix D for an example). At time 2, they were given five minutes to look over the corrections on their writing. Then their written texts were collected in by the researcher and they were given the same picture sheet and a blank writing sheet and told to rewrite the story (i.e., they were not able to consult their original story with the corrections). When they had completed the rewriting, their stories and the materials were collected by the researcher and they were

immediately given a new picture sheet and asked to write a new story. They completed the third writing task two weeks later.

Metalinguistic explanation group. The metalinguistic explanation (ME) group did not receive any individual corrective feedback. At time 2, they were given the metalinguistic explanation handout. The researcher read out the explanation to the class which took about five minutes and then gave the students five minutes to check their errors in their first story. The written texts and the ME handout were collected in by the researcher and the students then rewrote the story (i.e., without access to their original story or to the ME handout). When they had finished, their stories were collected and they were given a new picture sheet and asked to write a new story. The third written narrative was completed two weeks later.

Control group. The control group followed the same procedures as the two experimental groups but did not receive any DCF or ME ^[6]. They wrote the original story (Time 1), were given five minutes to prepare to rewrite it, rewrote it (without access to their original writing) and then wrote a second story (Time 2) and the final story at Time 3. Thus, the control group received the same practice in writing stories as the experimental groups but without any kind of feedback. In this way, it was possible to investigate whether there was any effect for writing practice alone as Truscott has suggested may occur.

Eye-tracking plus stimulated recall

Six students (three for the DCF treatment and three for the ME treatment) from another class of the same level (i.e., one not involved in the main study) participated in this part of the study. The purpose was to obtain information about how learners similar to those in the main study responded to the DCF and the ME 1) attended to the corrections (in the case of the DCF group) or their indefinite article errors (in the case of the ME group), 2) understood the rule for the use of indefinite article, and 3) used the rule to improve their text when revising it. The six learners first filled in the background questionnaire and wrote a story based on one of the picture compositions. They also completed the Error Correction Test as a pre- and as a post-test. These students did not complete the second and third writing tasks because of the potential reactive effect of the stimulated recall described below.

The scripts of three students in the DCF group were then corrected in the same way as for the other students. Two days later these three students were asked to read their story with the corrections on the monitor of an eye-tracking machine. Their stories, including the errors and the corrections, were transcribed in Microsoft Word Format in order to ensure the eye-tracking machine detected their eye-gaze movements. They were given five minutes to read their

corrected story, and then asked to rewrite the same story by hand without referring to their original script. In the stimulated recall, which was audio-recorded, the students were asked to watch a video recording of their eye-gaze movement while they read over their original piece of writing. The DCF learners recalled how they responded to the corrections in their revised text.

The ME participants received the ME handout at the beginning of the second session. The researcher read out the explanation to the participants and then asked them to read a version of their original script (with no corrections) on the monitor screen attached to the eye tracking machine. They were also given five minutes to check for errors. The ME sheet was placed on the desk between the monitor screen and the participants. Then the participants were asked to rewrite the same story by hand. In the stimulated recall, the ME students commented on any errors they had focused on during eye-tracking and also commented on what they had focused on when they rewrote the story.

Scoring and analysis

We think the statistical analyses are competent and hence we only have some minor comments:

1) Table 2 - What does asterix mean? We don't understand the rows labeled just "mean" - mean of what? Please explain. Standard deviations are needed for all means and they can be presented in brackets after the mean.

2) Probabilities currently presented as “p = .000” should be presented as “p < .001”

3) p15 - Results for error correction test - We think your post hoc comparisons also need to include pre-post (within group) comparisons. You've done this in later sections but not here.

Each indefinite article error in the Error Correction Test that was successfully corrected by a learner was awarded one point. The maximum score possible was six.

Obligatory occasion analysis was used to score the accuracy of each learner's use of the indefinite article in their written stories. An obligatory occasion was defined as any singular noun that was represented in the pictures with a referent that was [+specific referent/ -hearer knowledge]. This arose when a person or object or place was mentioned for the first time (e.g., when the student mentioned the “table” and “saw” in the task in Appendix A). The indefinite

article also occurs when the noun phrase is [-specific referent/ +hearer knowledge] (e.g., *A table is a kind of furniture*), or when it is [-specific referent/ -hearer knowledge] (*I don't have a table*). However, as these functions were not targeted in the study and not treated in the ME (see Appendix B) they were not included in the analysis. Both correct suppliance of the indefinite article in obligatory occasions and incidences of overuse were counted. Accuracy was calculated using Pica's (1994) formula for target-like use:

$$\frac{\text{Number of grammatical morphemes supplied accurately}}{\text{Number of obligatory contexts} + \text{Number of overused forms}} \times 100$$

To examine the reliability of the scoring 25 stories (over 15% of the total 44 stories) were re-scored by a second researcher. The overall agreement rate between the two scorers was 98%.

The scores for the Error Correction Test and the writing tasks were subjected to a series of repeated-measures ANOVAs^[7]. Pairwise comparisons using Bonferroni adjustment were computed where appropriate. Effect sizes for the comparative effects of two variables were estimated using the correlation coefficient *r*. Effect size values were interpreted as small = .1, medium = .3 or large = .5 as Cohen (1988) suggested.

The six participants' eye-gaze was recorded and analyzed using eye-tracker software, Tobii Studio 2.2. The eye-gaze data were analysed in terms of (1) the frequency of eye-fixations and (2) the duration of each eye-fixation. The "interest areas" for this analysis consisted of incidences of the use of or errors in the indefinite article. All the interest areas were set to a 2.0 x 3.5 cm rectangle. This information was then referred to in the stimulated recall sessions. The protocols arising out of the stimulated recall were transcribed and analyzed qualitatively to ascertain what use the learners had made of the DCF and ME when they read their corrected text and revised it.

Results

The results of the Error Correction Test and the accuracy scores for the written stories provide answers to research questions 1–3 (i.e., the effect of the DCF, the effect of the ME, and the comparative effect of the DCF and the ME). Descriptive statistics for the Error Correction Test and the target-like use scores for the four writing tasks are reported first followed by the results of the inferential statistical analyses. Finally, the results obtained from the eye-tracking, the stimulated recall and the Exit Questionnaire are reported to provide an answer to research question 4.

Descriptive statistics

The descriptive statistics for the Error Correction Test are shown in Table 1.

[INSERT TABLE 1 HERE]

Table 2 gives the descriptive statistics for the writing tasks. The scores indicate the proportion of correct use of the indefinite article as determined by target-like use analysis.

[INSERT TABLE 2 HERE]

Results for the Error Correction Test

A repeated-measures ANOVA showed that although there was no significant group effect ($F = 2.27$ (2, 40), $p = .177$), there was both a significant time effect ($F = 11.47$ (1, 40), $p = .002$) and a time-group interaction effect ($F = 10.08$ (2, 40), $p = .003$). The post-hoc pair-wise within group comparisons using the Bonferroni adjustment showed that only the ME group significantly improved from pre-test to post-test ($p = .000$) with a medium effect size ($r = 0.39$). Pairwise between-group comparisons (also using the Bonferroni adjustment) showed that there were no significant differences between the three groups in the pre-test. In the post-test the ME group significantly outperformed the control group ($p = .013$) with a large effect size ($r = 0.51$). The post-test differences between both the ME and the DCF and between the DCF and the control group did not reach statistical significance and the effect sizes were both small ($r = 0.26$ and 0.29).

Results for the accuracy scores in the written stories

We first examine the effects of the DCF and ME on the learners' revised texts. A repeated-measures ANOVA showed there were significant effects for time ($F = 14.11$ (3, 40), $p = .001$) and group ($F = 3.72$ (2, 40), $p = .043$) and also a significant time-group effect ($F = 6.40$ (6, 40), $p = .033$). The post-hoc pair-wise within group comparisons revealed that only the ME group significantly improved in accuracy of use of the indefinite article in the revised story ($p = .000$) with a large effect size ($r = 0.65$). In the between group comparisons, the ME group was found to significantly outperform the control group in accurate use of indefinite article ($p = .001$) with a large effect size ($r = 0.58$). Also the difference between the ME group and the DCF group approached significance ($p = .057$, $r = 0.43$). Although the DCF group also improved in accuracy in the revised story the difference did not reach statistical significance ($p = .534$, $r = 0.26$).

The groups' target-like use scores for the three stories were also compared. A repeated measures ANOVA showed that although there was no significant group effect ($F = 1.79 (2, 41), p = .180$), there was both a significant time effect ($F = 3.73 (2, 82), p = .028$) and a significant time-group interaction effect ($F = 3.20 (4, 82), p = .017$). The post-hoc pair-wise comparisons showed that the ME group improved significantly from Time 1 to Time 2 ($p = .002, r = 0.47$). However, this group's score declined at Time 3 and, as a result, the difference between Time 1 and Time 3 was no longer statistically significant ($p = 1.000, r = 0.07$). Neither the DCF group nor the control group showed any significant improvement over time. The only group difference to reach statistical significance was that between the ME group and the control group at Time 2 ($p = .003, r = 0.53$).

Learners' use of the DCF and ME

Table 3 shows the Error Correction Test scores of the six learners who completed the eye-tracking and the interview. It also shows the number of indefinite article errors they made in Story 1 and the corrections they made in the revised version of this story. The Error Correction Test scores in the pre-test ranged from 0 to 2 out of 6, indicating that all the learners failed to use the indefinite article with total accuracy. Two of the learners who received the ME (Maru and Morri) and one learner who received DCF (Medam) were successful in correcting their indefinite article errors when they revised their stories while the others (one who received the ME and two received the DCF) failed to correct their errors in the revised story.

[INSERT TABLE 3 HERE]

Eye-gaze movement

Table 4 shows the frequency and duration of the eye-fixations for each indefinite article error. There is no clear difference between the two groups in terms of the mean frequency of fixations or total duration of fixations. The highlighted parts in Table 4 indicate that the particular error was corrected by the learners when they revised their text. Hesah (ME), Arin (DCF) and Daria (DCF) failed to correct their errors even though they clearly attended to their original errors as shown by the number and duration of their fixations. On the other hand Morri (ME) and Medam (DCF) successfully corrected their errors in the revision even though their fixations were relatively infrequent and of short duration.

[INSERT TABLE 4 HERE]

Stimulated recall

All three participants who received the ME were able to explain the rule. Two of them (Morri and Hesah) stated that they did not know the rule before the treatment, indicating that the ME was successful in enabling them to understand it. Two of the three learners in the ME group were able to use the explanation they had been given to identify their errors in their original text. Maru, for example, demonstrated her ability to use the ME to identify an error:

Researcher: Did you find any errors here?

Maru: Yeah. I think this, “*a* rectangular table”.

R: Okay, this. You found any other mistakes?

M: No.

Hesah, however, explained that he had focused on identifying other types of errors in his text (e.g., third person -s). In other words, he had made no use of the ME sheet.

When it came to revising their original text, the three ME learners demonstrated different behaviors. Maru stated that she was very conscious about the rule she had learned. She used it successfully when revising. Morri also tried to make use of the rule. However, she sometimes overgeneralized the use of “*a*” to a context that required the definite article. Hesah also had an unclear understanding of the rule. For example, he stated that he only applied the rule to human nouns (e.g., *boy*) but not to non-human nouns (e.g., *table*). All of the three participants mentioned that they also tried to make their text more comprehensible. In other words, they did not focus exclusively on correcting article errors.

The participants who received the DCF experienced difficulty in working out the rule from the corrections they received. Daria did not know the rule and could not figure it out. She thought that “*a*” and “*the*” were interchangeable. Arin simply said that “*a*” should be used before nouns and adjectives. Medam did not pay attention to the corrections and stated that he did not think of the rule when he looked at his corrections:

Researcher: When you were looking at it, did you notice this [i.e., the error in “play on tennis table”]? Did you think “this is the first time, so I needed *a*”?

Medam: No.

R: You didn’t think about it?

M: No

His main strategy for dealing with the corrections was to try to memorize the complete text.

When they revised their texts, the DCF learners focused more on the content of their writing than on form. For example, Daria said, “I just think about these paragraph. I think I wrote I couldn’t write it very carefully”. They did not think about the grammatical rule for the use of the

indefinite article and were unable to explain it clearly. The only learner (Medam) who was successful in correcting his indefinite article errors did so intuitively. For example, asked why he correctly inserted “a” before the noun “saw” in his revised text he replied “I don’t know. I feel it sounds right.”

Table 5 provides a summary of the main findings of the interview. The ME learners developed an understanding of the rule for the use of the indefinite article although their understanding was sometimes idiosyncratic. They were able to use it to identify and reflect on their article errors. Two of the ME learners made use of the rule when revising their text. In contrast, the DCF learners failed to induce the rule from the corrections on their text and consequently were unable to use it when revising.

[INSERT TABLE 5 HERE]

Exit questionnaire

In their response to the multiple choice question, the majority of the learners in all three groups indicated that they thought the purpose of the lessons was “practicing writing” or “practicing general English skills”. In response to the open question about what they thought they had learned, two of the participants in the DCF and three in the ME group mentioned that they had learned how to use articles. Whereas the two DCF learners had prior knowledge of the rule as shown by the Error Correction Test (they scored 4 out of 6), none of the ME learners had done so, suggesting that what they had learned was due to the treatment. However, none of the other learners mentioned articles in their response to this question.

Discussion

The study had two main purposes – to investigate the effect of written error feedback on the development of learners’ explicit and implicit L2 knowledge and to compare one type of feedback (focused direct corrective feedback) that has been investigated in a number of previous studies and another type of feedback (metalinguistic explanation) that has not yet been investigated. The first purpose was theoretically motivated. While there is now evidence to show that written feedback leads to improved accuracy in new writing it is not clear whether this is because of the effect it has on implicit or explicit knowledge. As Polio (2012) pointed out, there is a need to establish an agenda for researching the effects of written error feedback on the development of learners’ explicit and implicit knowledge. The study we have reported was designed to make a start on this agenda. The second purpose was pedagogically motivated. We

wished to experiment with a type of feedback (metalinguistic explanation) that is less demanding on teachers' time than direct corrective feedback.

The first research question addressed whether the DCF had any effect on learners' acquisition of explicit and implicit knowledge of the English indefinite article. The DCF had no effect on scores in the Error Correction Test or on accuracy in new pieces of writing. Thus, the results of this study suggest that it had no effect on either their explicit or their implicit L2 knowledge. It should also be noted that the DCF did not lead to improved accuracy in the revised version of the first text.

The results of this study differ from those of previous studies. Ellis *et al* (2008) reported that that DCF led to statistically significantly higher scores in the same Error Correction Test that was used in this study and that the DCF group also outperformed a control group in this test. There are a number of possible reasons for difference in the results of these two studies. Whereas Ellis *et al* reported combined scores for indefinite and definite article, our study only investigated the indefinite article which research has shown to be later acquired than the definite article. Also, the learners in the two studies were very different. The Japanese university students in Ellis *et al* were experienced in learning explicit facts about grammar whereas the mixed learners in our study appeared less adept at handling metalinguistic information and, as the results of the Exit Questionnaire showed, treated the writing tasks as opportunities for practising writing rather than for learning grammar.

Other studies (e.g., Bitchener & Knoch, 2010; Ellis *et al*, 2008) reported a positive effect for focused DCF on both the accuracy of article usage in revision and in new pieces of writing. Again, there are a number of differences between these studies and our study. As in Ellis *et al*, Bitchener & Knoch examined accuracy in the use of both indefinite and definite articles while our study only investigated the indefinite article. One reason why the DCF was not effective in our study may have been because the corrective feedback was provided on only a single piece of writing in our study and several students received only one or two corrections. In both of the other studies several pieces of writing were corrected. Also, in Bitchener and Knoch (2010) the learners received multiple corrections. However, in Ellis *et al* (2008) many of the learners also only received one or two corrections of the indefinite article. Thus, the more likely explanation for the different results is the difference in the learners noted above.

The results of the eye-tracking (see Table 4) indicate that the three DCF learners who completed this did attend to the corrections. All three learners fixated multiple times on the errors. However, the overall duration of these fixations was relatively short suggesting that they may not have spent time reflecting on the nature of the errors. The stimulated recall data also suggest that these three learners attended to the corrections but revealed that they were not able to form an understanding of the rule for the use of indefinite article and use this understanding to

correct their errors when revising their text. The one learner who was successful did so intuitively rather than by applying the rule. Assuming that these three learners' eye fixations and responses in the stimulated recall responses are representative of the DCF experimental group, it appears that the DCF was ineffective because it failed to develop the learners' awareness of the rule. In other words, the learners may have been conscious of their errors at the level of "noticing" (Schmidt, 1993) but they were not conscious at the level of understanding and thus did not develop explicit knowledge. Clearly, too the DCF and the opportunity to revise were insufficient to develop implicit knowledge.

The second research question asked whether the Metalinguistic Explanation was effective in developing the learners' explicit and implicit knowledge. The ME group achieved significantly higher scores in the second administration of the Error Correction Test. Also, the ME group outperformed the control group in the post-test. This suggests that they were successful in improving their explicit knowledge of the indefinite article rule and were able to use this knowledge when revising.

The ME was also led to increased accuracy in a new piece of writing. However, the ME was not effective in the long-term. In the third story, written two weeks later, there was no statistically significant difference between the first and last pieces of writing and the effect size was minimal. In other words, the ME had a positive effect on immediate learning but its effect wore off over time. This suggests that although the ME assisted the development of the learners' explicit knowledge it had no effect on the development of their implicit knowledge. Implicit knowledge once developed is not easily forgotten so if the ME had had an effect on the learners' "genuine knowledge of language" the effect should have been durable. The results of the Error Correction Test administered after the learners had completed the third writing task showed that that the ME learners had improved their explicit knowledge. However, while they made use of their explicit knowledge in the second writing task immediately following the ME and the revision of their first piece of writing they do not appear to have done so in the third writing task perhaps because they were more focused on narrating the story (i.e., they were more focused on meaning rather than on form) and thus did not monitor their use of the indefinite article. As Krashen (1981) suggested explicit knowledge is only of value if learners are focused on form.

The eye-tracking and stimulated recall results indicate that the ME did lead to the learners attending to their errors in the first piece of writing and making use of the ME when revising. Maru, for example, only made one error in the use of the indefinite article, but fixated on this a total of 19 times for a total duration of 4.11 seconds. In the stimulated recall Maru demonstrated that he used his explicit knowledge to identify and correct his error. However, again what seems to have been important is not the "noticing" of errors by the ME learners but their "understanding" of the rule and their ability to use this to correct their errors. Hesah, for example,

fixated even more frequently than Maru on the errors he made and for a longer overall duration but admitted in the stimulated recall that he had paid little attention to the ME and focused on error types when revising

The third research question asked whether there was any difference in the effect of direct focused corrective feedback (DCF) and metalinguistic explanation (ME) on the development of the learners' implicit and explicit knowledge of the indefinite article. In line with the foregoing discussion, the results suggest that there was a difference for explicit knowledge but not for implicit knowledge. Only the ME group improved in the Error Correction Test and although it also improved in the second writing task (which functioned as the immediate test) this improvement was not sustained in the third writing task (which functioned as the delayed test). The DCF group showed no improvement in either the Error Correction Test or in the second and third writing tasks. The results suggest, then, that if the goal of written error feedback is to develop learners' explicit knowledge, ME may be a more effective means of achieving this than DCF.

The final research question addressed the use that the individual learners made of the DCF and the ME when revising their texts. To the best of our knowledge no previous study of written error feedback has examined this. Previous studies have focused exclusively on the outcomes of the feedback. In this study, we used eye-tracking and stimulated recall to investigate learners similar to those included in the main study. The eye-tracking results indicated that there was no difference in the attention that the DCF and ME learners paid to their correction/ errors. Both groups of learners fixated regularly on the relevant "interest areas" and for similar lengths of time. In other words, both groups engaged in "noticing". The stimulated recall results, however, suggest that there were differences in the "understanding" of the rule between the DCF and ME learners. The DCF learners clearly experienced difficulty in working out the rule for indefinite article and when they revised their texts they focused on content rather than form. In other words, they made little use of the feedback they had received. In contrast, the ME learners were able to explain the rule and tried to use it when they revised.

Conclusion

In the introduction to this article, we noted that L2 writing researchers (e.g., Polio, 2012; Bitchener, 2012) have suggested that written error feedback is likely to have an effect on L2 explicit rather than L2 implicit knowledge. Truscott (1996), too, argued that error correction may have an effect on learners' explicit knowledge but will not contribute to "genuine knowledge of language" (i.e., implicit knowledge). The study we have reported lends support to this claim. There was evidence that one of the types of feedback we investigated (metalinguistic

explanation) was effective in developing learners' explicit knowledge. For this to occur, however, the feedback had to result not just in learners attending to the errors they had made but to an understanding of the rule explaining the errors. Also, we concluded that feedback only had an effect on accuracy in new writing if learners used their understanding of the rule to focus on form and to monitor for grammatical correctness. While this appeared to have happened in the new writing produced shortly after the learners had received the metalinguistic information it did not occur in writing completed one week later. However, other types of learners (such as the analytical-type learners investigated in Ellis *et al*, 2008) may be better equipped to make more durable use of the explicit knowledge they have gained from feedback.

The results of the study suggest that neither the DCF nor the ME had any effect on learners' implicit knowledge. We argued that if the ME feedback had had an effect on implicit knowledge this would have been evident in the delayed as well as the immediate writing task but no such effect was evident. However, it would clearly be premature to rule out the possibility that feedback can impact on implicit knowledge as claimed by the weak-interface position and skill-learning theory (see Introduction). The DCF in our study involved very few corrections for some individual learners. It is possible that more intensive feedback combined with the "pushed output" that revision provides will facilitate the development of implicit knowledge over time. The ME was also quite limited – it was a one-shot treatment. It is possible that reminding learners of the rule prior to subsequent writing tasks will help them to proceduralize it.

To further the "agenda" that Polio argued is needed, further thought needs to be given to how implicit and explicit knowledge can be measured. We believe that the Error Correction Test provided a valid measure of explicit knowledge. However, writing tasks – even when pressured and learners are primarily focused on meaning – cannot afford a convincing measure of implicit knowledge given that they allow for controlled processing and monitoring for accuracy. In future studies, therefore, it may be necessary to make use of the kinds of instruments that SLA researchers have employed to measure implicit knowledge (e.g., oral elicited imitation or free oral production). The difficulty here, of course, is that such instruments may lack face validity given that the aim of written error feedback is to help learners improve their writing.

The results of the study showed that for the low-proficiency learners we investigated, ME was more effective than DCF in developing learners' explicit knowledge. DCF was not sufficient to enable the low-proficiency learners to form an understanding of the target rule and did not help them to either revise or write a new story. In contrast, the ME treatment was successful. It enabled learners to develop their explicit knowledge of the indefinite article rule and to make use of this knowledge when revising and writing a new story. However, as already noted, its effect on improving accuracy in new writing was not sustained. The study suggests that if the goal is to develop explicit knowledge and encourage monitoring, ME constitutes an effective and practical

means of doing so but that it may be necessary to remind learners of the rule they have been taught prior to new writing tasks.

A feature of our study is that we investigated how learners responded to the feedback, using eye-tracking and stimulated recall. The information that these provided proved invaluable in helping to see if learners had attended to the corrections, had developed an understanding of the target rule, and had used it when revising. Future studies of written error treatment might like to employ these and other methods to examine how learners process and use feedback.

Acknowledgments

We would like to the reviewers of earlier drafts of this article. We are grateful to the faculty members and students at the English Language Centre in Michigan State University for supporting this study. In particular, we are grateful to Jennifer Behney for helping us to organize the data collection and Sehoon Jung for help with the eye-tracking.

Commented [JSLW4]: Would you also consider thanking the JSLW reviewers? Entirely up to you.

Notes

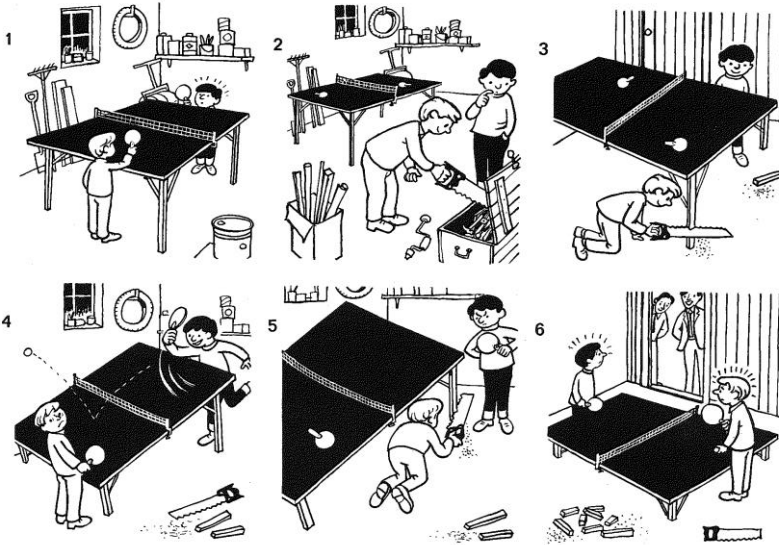
1. Frantzen (1995) investigated the dual effect of explicit grammar instruction and indirect CF on accuracy in a grammar test and in writing. Frantzen reported an effect in the case of the grammar test but not in writing. The study reported in this article differed from Frantzen's study in that it sought to compare the effect of CF and explicit instruction.
2. The fact that errors corrected in one piece of writing do not occur in a subsequent piece of writing provides an explanation for why some studies that have investigated unfocused CF have failed to find any effect for CF. Bruton (2009) pointed out that this may have been the case in Truscott and Hsu's (2008) study.
3. A reviewer of a draft version of this article claimed that focused CF studies are flawed as "they don't look at what else is happening in writing". While we accept that focused studies do "present a narrow picture of what is happening" we see this as a strength in experimental research where the purpose is to establish whether CF directed at a single feature has any effect on learners' implicit and explicit knowledge. Other studies can look at the wider picture. We also consider focused CF to have ecological validity as explained in the introduction.
4. However, the exit questionnaire administered to students in Ellis *et al.*'s (2008) study showed that the students did not treat the focused task as a "written grammar exercise" but as an opportunity to practice writing and develop their general L2 proficiency. Indeed, given that the majority of focused CF studies have investigated English articles, which are ubiquitous in any text, it is unlikely that students performing these tasks perceive of them as requiring a

primary focus on grammar.

5. A reviewer felt that a two week gap was insufficient for the delayed test. Unfortunately, for logistic reasons a longer gap was not possible. Also, the positive effect for ME in the second piece of writing was largely lost in the third so, in fact, a two week gap proved sufficient to establish whether or not the effect was sustained.
6. The unequal treatments between the three groups were compensated for by providing DCF and the ME sheet after the project to those participants who requested it.
7. As some of the data failed to meet normality assumptions (the ME group and the Control group in Time 1 and the Control group in Time 3), these data were also analyzed with non-parametric tests using ranked data. The results were equivalent to those obtained from the ANOVA.

Appendix A. An Example of the Picture Composition

The table that got smaller



Appendix B. Error Correction test

Name _____ Date / /

Error correction task

Please read each statement. Each statement has two sentences that are related. One of the sentences is underlined. The underlined sentence contains one error. Write out the underlined sentence correcting the error. (Note: There are no punctuation or spelling errors.)

Example 1. Gloria has lived in Auckland since 2010. She really enjoys living in Auckland.

Answer: Gloria has lived in Auckland since 2010.

Example 2. John got a cold. He couldn't went to school yesterday.

Answer: He couldn't go to school yesterday.

1. Mary used to living in Chicago. She lives in New York now.

2. I look after a little girl on Saturday. Little girl is very smart.

3. The driver's tyres were in bad condition. They were caused a bad accident.

4. I read book about New York. The author, however, was from California.

5. I took a test yesterday. I couldn't finish test.

6. I saw a man in car across the street. I realized that he was my brother.

7. I saw a very interesting movie last night. I have forgotten the name of movie.

8. Last night I read a newspaper. I don't know where newspaper is today.

9. A young woman and a tall man were talking outside my house. Then the woman started shouting at man.

10. We rented boat last summer. Unfortunately, the boat hit another boat and sank.

11. We watched movie on Saturday. The actors in the movie were all very famous.

12. When you turn onto Paramus Road, you will see two houses: a blue one and a yellow one. I live in blue house.

13. Is your uncle car salesman? I'm looking to buy a car.

14. Bill was so drunk last night. He couldn't even recognized his girlfriend.

15. We bought a new car last week. The car stolen yesterday.

16. My mother was fired yesterday. She will have to find new job.





-1-





-2-

Appendix C. Metalinguistic Explanation

How to use 'a (an)' and 'the'

Using 'a (an)' and 'the' is difficult, but there is one helpful rule. When you tell a story, you often repeat the same words. For example:





One day,  was practicing with some   came to the park and started to hit .





 was surprised and asked  to stop hitting them. But  ignored her, so  was very angry.

Now let's see how to use 'a (an)' and 'the'?

1. When you mention an object for the first time, you use 'a (an)' before the noun (*golfer player, baseball player*). If the object is more than one (*golf balls*) and you mention them for the first time, you don't need any article ('a(an)' or 'the').
2. Whenever you mention the same object/ objects again, you use 'the' before the noun.

So the answers are:

One day,  was practicing with some   came to the park and started to hit .

 was surprised and asked  to stop hitting them. But  ignored her, so  was very angry.

So the rule is:

First time	One thing → 'a (an)'	More than one → NO article
Again	'the'	

➤ Now please check errors in your writing. Later, you will re-write your story without seeing the original story or this sheet.

Appendix D. An Example of Direct Written Corrective Feedback

Please write a story based on the pictures you have been given. The story should be detailed enough to understand the story without seeing the pictures. Use your imagination to fill out details in the story. You have 20 minutes to write.

There are Family from mother, father, and two sons. The father and mother went out and they said for their sons to stay at home and play without any accident and the sons said, "ok, but when you come back please bring with you candy and chips, ok". The mother and father say "OK, be quiet and carefully". The two sons went to play by a table game, but it was too high for them. Then they decided to cut the legs of the table from one side. After that, they started to play, but they faced a big problem which is one of them became always winner because the lower side of the table help him. The other boy was very angry and he didn't like that. Then he decided to cut the other leg of the table. Then the table was good for them and they started playing. Finally, the father and mother came back to their house, and they are surprising; the father just laughing and the mother was confused "what happened here?".

References

- Andersen, T. (2010). *The effect of tiered corrective feedback on second language academic writing*. Master of Arts thesis, University of British Columbia, Vancouver. Retrieved from <https://circle.ubc.ca>
- Bitchener, J. (2012). A reflection on ‘the language learning potential’ of written CF. *Journal of Second Language Writing*, 21(4), 348-363.
- Bitchener, J. & Ferris, D. (2011). *Written Corrective Feedback in Second language Acquisition and Writing*. Place: Routledge.
- Bitchener, J., & Knoch, U. (2008). The receptiveness of international and migrant ESL students to written corrective feedback. *Language Teaching Research Journal*, 12 (2), 409-431.
- Bitchener, J. & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31(2), 193-214.
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207-217.
- Bruton, A. (2009). Designing research into the effect of error correction in L2 writing: Not so straightforward. *Journal of Second language Writing* 18(2), 126-140.
- 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.
- DeKeyser, R. (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*. Cambridge: Cambridge University Press.
- Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.
- Egi, T. (2007). ‘Interpreting recasts as linguistic evidence: the roles of linguistic target, length, and degree of change’. *Studies in Second Language Acquisition*, 29(4), 511-37.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352
- Ellis, R. (1997). *SLA research and language teaching*. Oxford ; New York: Oxford University Press.
- Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, J. Philp & H. Reinders. *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Bristol: Multilingual Matters.
- 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* 26, 353-371.
- Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language*

Commented [JSLW5]: For consistency, please add issue number, here and in all journal articles, please

acquisition research. London: Routledge.

- Farrokhi, F. & Sattarpour, S. (2012). The effects of direct written corrective feedback on improvement of grammatical accuracy of high-proficient L2 learners. *World Journal of Education*, 2(2), 49-57.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes. A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1-10.
- Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge: Cambridge University Press.
- Ferris, D. & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10, 161-84.
- Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28, 323-51.
- Fotos, S. & R. Ellis. (1991). Communicating about grammar: a task-based approach. *TESOL Quarterly*, 25, 605-28.
- Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate content course. *Modern language Journal*, 79(3), 329-344.
- Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, 50, 343-84.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* San Diego: Academic Press.
- Mohamed, N. (2001). *Teaching grammar through consciousness-raising tasks*. Unpublished MA thesis, University of Auckland, Auckland.
- Ortega, L. (2012). Epilogue: Exploring L2 writing–SLA interfaces. *Journal of Second Language Writing*, 21(4), 404-415.
- Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21(4), 375-389.
- Robb, T., Ross, S., & Shortreed, I. A. N. (1986). Salience of feedback on error and Its effect on EFL writing quality. *TESOL Quarterly*, 20(1), 83-96.
- Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29, 67-100.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics* 13, 206-226.

- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15(2), 147-163.
- 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.
- Truscott, J. & Hsu, A. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 7(4), 292-305
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-69.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research* 14, 103-135.
- Truscott, J. (1999). The case for “ the case for grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8, 111-22.
- Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, 13, 337-343.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing* 16, 255-272.
- Truscott, J. (2010). Some thoughts on Anthony Bruton's critique of the correction debate. *System*, 38(2), 329-335.
- Van Beuningen, C. G., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1-41.
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321-331.