Understanding Written Corrective Feedback in Second-Language Grammar Acquisition

Jason Paul Wagner[1] & Douglas J. Wulf[1]

[1] Department of Linguistics, George Mason University, Fairfax, Virginia, USA

Correspondence: Jason Paul Wagner, Department of Linguistics, George Mason University, Fairfax, Virginia, USA. E-mail: wagnerjp@aol.com

Abstract

Written Corrective Feedback (WCF) is used extensively in second-language (L2) writing classrooms despite controversy over its effectiveness. This study examines indirect WCF, an instructional procedure that flags L2 students' errors with editing symbols that guide their corrections. WCF practitioners assume that this guidance will lead to increased grammatical competence over time in new writing samples. This study finds that these assumptions are correct overall. However, in-depth analyses of L2-English learners' correction behaviors in four elicitation tasks over a 12-week period demonstrate that WCF is not uniformly effective at increasing accuracy for all grammatical constructions. In fact, WCF fails to exert any positive effect with a number of grammatical constructions. This result can be understood via Skill Acquisition Theory (SAT) when the treatability of constructions with WCF is considered. Specifically, grammatical constructions that include only a binary option for correct usage are highly amenable to positive change via WCF since employing WCF is akin to correcting errors flagged on a true/false test. By contrast, grammatical constructions with more than a binary choice for correct usage, akin to correcting errors flagged on a multiple-choice test, are not amenable to positive change.

Keywords: English grammar, English as a second language, second language acquisition, skill acquisition theory, written corrective feedback

1. Introduction

Using Written Corrective Feedback (WCF) to teach grammar in second-language (L2) writing classrooms is widespread despite controversy about whether it produces more accurate L2-student writers. Indirect WCF is an instructional procedure whereby writing errors are not directly corrected, but rather just flagged with proofreading marks (e.g., *s-v* for a subject-verb agreement error, *vt* for a verb tense error, etc.). Students are then challenged to correct these flagged errors in subsequent drafts of the same writing sample. It is hoped that, over time, revision via WCF will lead to fewer errors (i.e., improved accuracy and grammatical competence) in new writing samples. Unless noted otherwise, in discussing WCF, we refer to this variety of indirect WCF and not to other kinds of writing feedback that supply information beyond commonly used WCF proofreading marks.

WCF research in Second-Language Acquisition (SLA) has been extensive, yet criticized for methodological flaws. Also, the research has arguably been too broad or too narrow in its scope. Broadly focused studies (Lalande, 1982; Robb, Ross, & Shortreed, 1986; Fathman & Whalley, 1990; Kepner, 1991; Truscott, 1996, 1999, 2004, 2007; Polio, Fleck, & Leder, 1998; Ferris, 1999, 2004, 2010; Ashwell, 2000; Ferris & Roberts, 2001; Truscott & Hsu, 2008) only consider whether WCF works overall without examining WCF's impact on specific grammatical constructions. Narrowly focused studies (Sheen, 2007; Bitchener, 2008; Bitchener & Knoch, 2008; Ellis, Sheen, Murakami, & Takashima, 2008; Bitchener & Knoch, 2009, 2010; Sheen, 2010; Shintani & Ellis, 2013) have considered only one grammatical construction or a part of a single grammatical paradigm, but such findings are not highly applicable to the L2 classroom and do not clarify whether WCF helps students acquire complete grammatical paradigms.

The aim of this empirical study is to scrutinize WCF as it is actually used in L2-writing classrooms, ensuring methodological controls and conducting statistical analyses that provide a detailed picture of WCF effectiveness. We examine not only overall effectiveness, but also effectiveness for different error categories (i.e., grammatical, mechanical, and word usage) and for specific grammatical constructions.

Our results indicate that constructions with a binary usage paradigm are amenable to improving accuracy via WCF over time. One such example is English singular/plural, which involves adding -*s* versus omitting -*s*, and which is only slightly complicated by a number of irregular plurals, such as *goose/geese*. Correcting errors in a binary paradigm is indeed as easy as correcting mistakes flagged in a true/false test. However, whenever the grammatical paradigm is non-binary (i.e., having more than two options), WCF is shown to be useless in that it does not contribute to improving accuracy over time. WCF with greater than binary paradigms is akin to flagging errors on a multiple-choice test. The feedback indicates that a given form is incorrect, but does not indicate what the correct form is. Therefore, WCF does not provide sufficient information to assist the learner in acquiring the correct forms of the paradigm in a positive way.

The current study applies Skill Acquisition Theory (SAT) (Anderson, 1976; McLaughlin, 1987; DeKeyser, 1998, 2001, 2007) to explain these results. SAT regards language acquisition as a specific instance of the more general phenomenon of skill acquisition. According to SAT, acquisition of any skill, such as learning to play a sport, typically begins when knowledge about the skill (declarative knowledge) is imparted to the learner (e.g., *"Hold the baseball bat all the way back, swing the bat straight across, and follow through."*). Then, the learner must implement the declarative knowledge through extensive practice in order to build procedural knowledge of that skill (e.g., swinging at baseballs in the batting cage extensively). As practice continues, the skill is developed, and the more one practices, the better the learner can become at the skill. Ultimately, procedural knowledge can become internalized such that it is automatic and unconscious (e.g., swinging the bat correctly and effectively without consciously thinking about the declarative knowledge). We see this developmental pattern of skill learning occur with athletes, musicians, skilled laborers, and, as has been noted by supporters of SAT, L2 learners.

Under the SAT paradigm, it is possible to realize how WCF may sometimes be effective and sometimes not. Simply stated, SAT maintains that WCF would be useful whenever it contributes sufficiently to the learner's declarative knowledge of grammatical constructions. The relevant declarative knowledge, in this case, is a conscious grasp of how the grammar works. Armed with this information, learners can then practice the forms by correcting them in subsequent drafts of the same writing sample. This practice acts as a bridge from the declarative knowledge imparted by WCF to procedural knowledge of these same forms. Here, the procedural knowledge is the ability to use the grammar accurately in actual language production, as in academic writing.

The concept of proceduralization is operationalized in the current study by noting accuracy changes of experimental group participants (who receive WCF) when compared to accuracy changes of control group participants (who do not receive WCF) over time. Changes in accuracy can be used to demonstrate that declarative knowledge plausibly has transitioned to proceduralized knowledge in the writing modality. Experimental participants who outperform control participants in accuracy over time demonstrate this transition. They reduce errors significantly on constructions even though they are unaware that their accuracy changes are being examined and analyzed.

Only in cases of binary paradigms does WCF impart a decipherable repair strategy that can be applied in subsequent writing samples. The correction process can then eventually lead to greater proficiency with the grammatical paradigm. For non-binary grammatical paradigms, WCF fails to impart the necessary information, rendering WCF ineffective even when used over an extended period of time.

2. Previous WCF Research

Before describing the present study, it is helpful to consider the nature of the ongoing debate over WCF. First of all, although WCF is a widely used pedagogical practice, its status among those researching its use remains controversial, with views ranging from strong support to total rejection. Ferris (1999) claims that there exists "mounting research evidence that effective error correction—that is selective, prioritized, and clear—can and does help at least some student writers" (p. 4). By contrast, Truscott (1996, p. 328) argues, "Grammar correction has no place in writing courses and should be abandoned", claiming that "research evidence shows that grammar correction is ineffective".

Part of the reason for such widely contrasting views, we suspect, may be due to unevenness in WCF effectiveness that has not been documented previously. This has not become apparent in part because many WCF studies have sought only to measure the overall effectiveness of WCF. That is, studies have simply tracked experimental groups receiving WCF to see if they exhibit fewer total errors in comparison with control groups over time. Yet, such broadly focused scholarship has not provided any clear determination of whether WCF is effective uniformly. In addition, there have been a number of studies with the very narrow focus of WCF's impact on acquiring some portion of the article system in English. Here again, this research focus cannot tell us

if WCF is a one-size-fits-all error reduction system, as one might be tempted to assume it to be. For this reason, our study takes a separate look at WCF's impact on each kind of error. Indeed, our findings reveal just such a lack of uniform effectiveness, a discovery, we feel, that is a significant contribution to the WCF debate.

Ferris (1999) touches briefly on the idea that WCF may not have uniform effectiveness across all error types or grammatical categories. She introduces the concept of *treatable errors* versus *non-treatable errors*, defining this distinction based on whether an error is *rule based* (i.e., there are rules that can be taught, learned, or looked up in a grammar book) or *idiosyncratic* (i.e., there are no systematic rules to learn). She argues that rule-based systems should be treatable via WCF, claiming success in her own classrooms using WCF to correct errors such as subject-verb agreement, run-ons and comma splices, missing articles, and verb form errors. On the other hand, she argues that idiosyncratic errors, including lexical errors and problems with sentence structure, should not be as amenable to change via WCF. Ferris' notion of treatable and non-treatable errors is a useful one, although the results of our study lead us to conclude that this contrast relates to whether a grammatical paradigm is binary or non-binary.

In addition, it should be noted that previous WCF studies have conspicuously met with criticism of two kinds: methodological flaws and lack of applicability to the language classroom. So, to begin with, the results of many of the broadly focused studies have been called into question due to their research methodology. Thus, Truscott (1996) examines Hendrickson (1978), Higgs (1979), Raimes (1983), Semke (1984), Robb et al. (1986), Fathman and Whalley (1990), Hyland (1990), Bartram and Walton (1991), Kepner (1991), and Sheppard (1992) noting lack of control groups, lack of statistical analysis of data, and the need for better measurement instruments (e.g., writing samples produced by the students rather than grammar exercises). Ferris (2004) examines thirty WCF studies and notes that only six examined an experimental group against a control group—Semke (1984), Fathman and Whalley (1990), Kepner (1991), Polio et al. (1998), Ashwell (2000), and Ferris and Roberts (2001). Of these, only two (Kepner, 1991; Polio et al., 1998) examined the effects of WCF over a significant time interval (i.e., weeks or months). Of the six studies with a control group, four provide support for WCF (Fathman & Whalley, 1990; Kepner, 1991; Ashwell, 2000; Ferris & Roberts, 2001), one is missing information and is inconclusive (Semke, 1984), and one reports no advantage for the group that received WCF (Polio et al., 1998). Consequently, only two studies, Kepner (1991) and Polio et al. (1998), seem entirely sound methodologically, but these reach opposite conclusions on WCF. In addition, Truscott (1996) interprets Kepner (1991) not to show support for WCF.

Indeed, criticism of WCF research methodology has led to reinterpretations of many studies (Truscott, 1999; Chandler, 2003; Ferris, 2004; Guenette, 2007; Bitchener & Knoch, 2008). In fact, so many reexaminations have occurred that some WCF researchers have complained that old studies are constantly reinterpreted at the expense of doing more methodologically sound WCF research, and Bruton (2010, p. 491) characterizes the debate as "tedious, sterile, and academic." Crucially, the present study has been designed to avoid the criticisms over methodology leveled at so many previous WCF studies. We feel our study methodology has yielded results that can be confidently relied upon as being informative.

By contrast, the narrowly focused studies of WCF and the article system (Sheen, 2007; Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010; Ellis et al., 2008; Sheen, 2010; Shintani & Ellis, 2013) have been criticized for yielding results that are uninformative for classroom instructors. Such studies have argued for WCF's effectiveness with improving two functional uses of the English article system: (1) use of the indefinite article when mentioning something for the first time, and (2) use of the definite article when the same thing is mentioned subsequently. Generally speaking, experimental group participants in these studies outperform control groups in correcting these two functional uses of English articles over time. However, such studies only demonstrate that a subset of the usage paradigm of the English article system is amenable to improvement. These studies can neither confirm nor refute whether WCF can aid in the acquisition of the entire English article system. As a result, the general applicability to L2-writing classrooms is questionable since teachers need to decide whether to treat article errors with WCF or not. However, we regard our study as providing insights that can be directly informative to language teachers who want to know whether to employ WCF and, if so, exactly how.

It should also be noted that the findings of studies of WCF and the article paradigm vary. Bitchener, Young and Cameron (2005) find improvement in the use of the definite article only. Ferris, Chaney, Komura, Roberts, and Mckee (2000) find that students who received WCF on article usage actually regressed in their abilities to produce articles correctly over the elicitation period. Although Ferris and Roberts (2001) find gains in the accurate use of articles overall compared to a control group, this study is not longitudinal in that only 2 weeks separated the first and final drafts examined. Also, this study only investigates WCF effectiveness as it relates to

first and final drafts of the same writing sample, and does not evaluate WCF as a teaching/learning tool that aids in learning and long-term acquisition. Thus, the literature on the topic of WCF effectiveness with article usage does not demonstrate that WCF is effective in improving use of the article system as a whole. Our results, in fact, also indicate that WCF is ineffective in helping learners improve with article usage overall.

3. The Study

3.1 Research Questions and Hypotheses

Our study asks the following research questions:

- RQ 1: Does learners' overall accuracy (reduction of all error types) improve over the course of the 12-week research period as a result of WCF?
- RQ 2: Does learners' overall grammatical, mechanical, and word usage accuracy improve over the course of the 12-week research period as a result of WCF?
- RQ 3: Are individual grammatical errors equally amenable to change with the use of WCF?
- RQ 4: Is WCF effectiveness comparable for low and high proficiency students?

For each research question, the proficiency difference of the experimental group and control group is judged statistically significant in a one-way repeated ANOVA statistical analysis with an F-Statistic <.05. An F-Statistic ≥.05 indicates the experimental group did not significantly outperform the control group.

3.2 Study Participants

Study participants originally consisted of 40 students from a high school in Northern Virginia. These were at intermediate and advanced English proficiency levels (21 intermediate; 19 advanced). Over the course of the 12-week research period, 7 participants were released from the study because they moved out of the area, attended too few classes to complete the elicitation tasks, or changed their class schedule at the school. As a result, the study includes 33 participants (n=33): 17 intermediate and 16 advanced participants.

In Virginia, students' initial English proficiency and their annual proficiency progress are measured using World Class Instructional Design and Assessment (WIDA) proficiency evaluations. WIDA assesses reading, writing, speaking, and listening skills as well as social and academic language skills. Composite scores (i.e., overall proficiency) are based on a combination of students' four individual language skill scores and their social and academic language scores. Students' overall proficiencies are ranked from 1 to 6 (wida.us/standards/CAN_DOs/, 2009).

Participants are predominately Central and South American immigrants. The majority are from El Salvador, but participants also come from Honduras, Guatemala, Vietnam, Laos, the Philippines, Romania, Bangladesh, Afghanistan, Egypt, and Iran. These ninth and tenth grade students' ages range from 14 to 18. Their grade levels are determined by a combination of previous education in their home countries, English proficiency, and the number of mainstream classes already completed in the U.S. Intermediate students generally have had at least one year of instruction in the U.S. and a WIDA proficiency score between 1.0 and 2.9. Intermediate students take very few mainstream classes, if any, as their class schedules include English Language Learner (ELL) sheltered classes. Advanced students generally have had at least two years of instruction in the U.S. and a WIDA proficiency score between 3.0 and 4.9. Most of these students take a majority of mainstream classes.

3.3 Study Design

3.3.1 Participant Placement

Based on WIDA scores, the school places students into one of two intermediate ELL writing classes or one of two advanced ELL writing classes. For this study, intermediate-level participant groups are designated as LOW and advanced-level participant groups are designated as HIGH. Flipping a coin determined which LOW and HIGH groups would serve as control groups (receiving no WCF) and experimental groups (receiving WCF). Table 1 summarizes participants' information and group placement.

Table 1. Participant and group information

Group	Number of Participants	Proficiency	Home Countries	Control/Experimental
1	9	LOW	Iran 1 El Salvador 5 Guatemala 1 Vietnam 1 Philippines 1	C-No WCF
2	8	LOW	Saudi Arabia 1 El Salvador 5 Philippines 1 Laos 1	E-WCF
3	8	HIGH	Egypt 1 El Salvador 5 Vietnam 1 Philippines 1	C-No WCF
4	8	HIGH	Egypt 2 El Salvador 3 Honduras 1 Romania 1 Vietnam 1	E-WCF

3.3.2 Instruction

Apart from the crucial WCF contrast, a concerted effort was made to hold all other variables among the four classes constant. The instructor for both experimental and control groups was the same individual (one of the authors of this paper). All classes received 4.5 hours of instruction per week. Over the 12-week treatment period, all instructional content having to do with writing was the same among all classes (i.e., writing paragraphs in different rhetorical forms). All classes began with writing individual journals that used the same prompts and that were never corrected for grammar, punctuation, or mechanics. WCF was reserved for academic paragraph writing only in the experimental groups.

Grammar, punctuation, and word usage mini-lessons were delivered in the same way in each class, except that the experimental groups received WCF correction symbol keys. Before elicitation tasks were undertaken, instruction was given to all participants (control and experimental) on each error category. These lessons were concluded before the 12-week elicitation task period began so as not to affect error rates during the elicitation period.

Each of the classes followed a five-step writing process for completing each elicitation paragraph. The five steps of the writing process are: (1) prewriting, (2) organizing, (3) drafting, (4) revising/editing, and (5) publishing (i.e., producing and submitting final drafts with corrections). However, only the experimental groups received WCF as a part of step 4. The experimental groups were then asked to publish their paragraphs using the correction symbols and the correction symbol key to make the proper corrections. The two control groups were simply asked to "look for errors and try to correct them" before and during the publishing stage.

3.3.3 Targeted Error Categories and Types

Eleven error types were targeted and tabulated. These error types can be further categorized by their intended instructional purpose, as shown in Table 2:

jel.ccsenet.org
Journal of Education and Learning
Vol. 5, No. 4; 2016

Table 2. Error categories and types

Error Category	Error Type
Linguistic/Grammatical (4)	Subject Verb Agreement/Verb Tense/Singular-Plural/Article Usage
Word Usage (2)	Wrong Word (used for homonyms only)/Word Form (parts of speech)
Mechanics (5)	Capitalization/Punctuation/Run-ons/Fragments/Spelling

Examining these error categories would indicate if participants could recognize and acquire particular linguistic forms while simultaneously attending to other error types. Table 3 provides the WCF correction symbols.

Table 3. Correction symbol key

Symbol	Symbol Meaning	Example
s-v	subject-verb agreement	He <u>love</u> his wife.
num	number- singular/plural	I have many <u>dog</u> at home.
cap	capitalization	<u>doug</u> went to NYC on <u>t</u>uesday.
ro	run on	<u>He likes to go to the store, and he likes to go to the beach, but he does not like to go to the park when it is hot because there is no place to cool down, and his sister also hates the park in the summer time, but their mother loves to walk in the park any time of the year, so she goes there every day to exercise</u>.
vt	verb tense	I <u>will be</u> in class yesterday.
frag	fragment	<u>If you like pizza</u>.
wf	word form	He completed his <u>apply</u> for college last night.
ww	wrong word	<u>There</u> house is very beautiful.
P	Punctuation (. ? ! , ; : ' ")	What is he doing<u>__</u>
sp	spelling	My <u>freind</u> is <u>funy</u>.
art/no art	wrong or missing article (a, an, the)	My teacher is <u>a</u> best one in ____ school

3.3.4 Elicitation Tasks

Each participant was tasked to complete four 200-word paragraphs in four rhetorical forms over the 12-week elicitation period at weeks 1, 4, 8, and 12. All participants were given explicit instruction in the four rhetorical forms: (1) Description, (2) Exemplification, (3) Compare/Contrast, and (4) Cause and Effect. After completing prewriting and organizing steps, students were given 45 minutes to complete paragraph drafts. Correction symbols were added to the experimental groups' paragraphs between classes, and all classes published their final drafts in the following class.

3.3.5 Statistical Analysis

ANOVA with repeated measures was used to examine error rate changes across the four elicitation tasks and differences between control and experimental groups. Errors were tallied from the first 150 words of each paragraph, since not every participant consistently attained the 200-word length requirement.

4. Study Results

4.1 Overall WCF Impact

The measure of WCF impact broadly for all error types (linguistic/grammatical, word usage, and mechanical) demonstrates a statistically significant difference in the reduction of errors over the four elicitation tasks between participants from both LOW and HIGH groups (ANOVA F-Statistic=.004), with experimental groups reducing errors significantly better than control groups over the elicitation period (i.e., a statistical significance level of <.05). Table 4 shows relevant statistics with Figure 1 providing a graphic representation of the same data.

Table 4. Total errors (Includes all Error Types) (Descriptive statistics for mean test scores by group and elicitation task)

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Control	17	26.12	10.222	24.71	7.622	21.41	8.163	25.29	9.116
Experimental	16	27.50	13.638	19.12	9.824	17.81	8.542	14.88	8.500
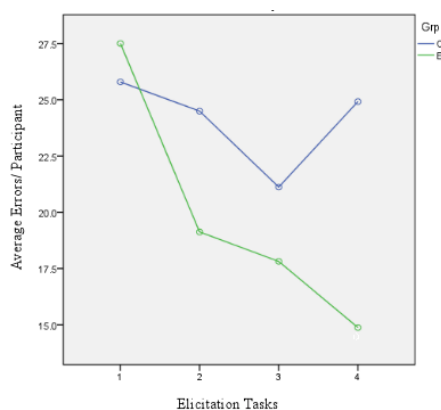
Figure 1. Effectiveness of WCF on total errors over time

Thus, the answer to RQ1 is yes.

4.2 WCF Impact on Specific Error Types

Specific grammatical errors tracked were subject-verb agreement, verb tense, singular/plural, and article usage. As the results indicate, experimental groups reduced grammatical errors as a category significantly better across the four elicitation tasks relative to the controls (ANOVA F-Statistic=.045), as shown in Table 5 and Figure 2. This means that participants were able to improve linguistic accuracy while attending to other types of corrective feedback.

Table 5. Grammatical errors (Subject-Verb, Verb Tense, Singular/Plural, and Articles) (Descriptive statistics for mean test scores by group and elicitation task)

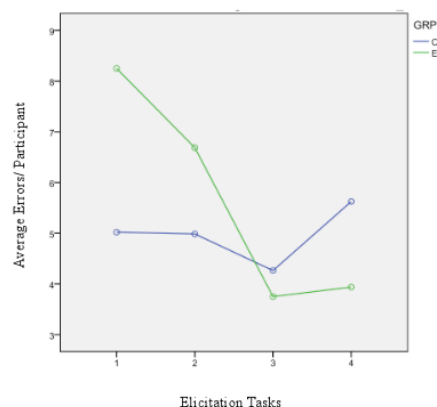| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	5.12	3.352	5.06	3.092	4.35	2.548	5.76	3.383
Experimentals	16	8.25	5.260	6.69	6.096	3.75	3.337	3.94	3.395

Figure 2. Effectiveness of WCF on linguistic errors over time

As seen here, experimental groups reduce their errors by over 50% from elicitation task one to four, whereas control error rates stay flat and even increase slightly in the final elicitation task.

Experimental groups also significantly outperformed controls over time on reducing word-usage errors (ANOVA F-Statistic=.044). In particular, wrong word errors involving homonym selection (symbol "ww" for homonyms) (ANOVA F-Statistic=.026) were effectively treated via WCF.

There was also a marginal statistical difference with control groups in reducing mechanical errors (ANOVA F-Statistic=.086, a statistical significance level of <.1). Of all mechanical errors, the data indicate that punctuation is most amenable to change (ANOVA F-Statistic=.036). A majority of punctuation errors were incorrect comma uses, which involved the binary paradigm of adding or omitting a comma.

Thus, considering performance in all four error categories, the answer to RQ2 is yes.

4.3 Individual Linguistic Errors and WCF Effectiveness

4.3.1 Subject-Verb Agreement and Singular/Plural

Of the four linguistic errors types examined, subject-verb agreement is most amenable to change. A statistically significant difference was discovered between the controls' and experimentals' ability to reduce subject-verb agreement errors over time (ANOVA F-Statistic=.008), as seen in Table 6 and Figure 3.

Table 6. Subject-verb agreement errors (Descriptive statistics for mean test scores by group and elicitation task)

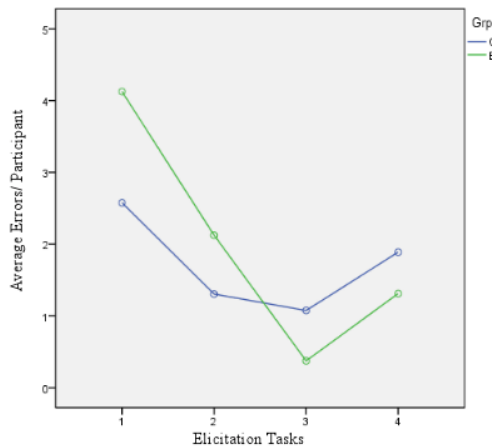| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	2.65	2.691	1.29	1.404	1.12	1.054	1.94	1.853
Experimentals	16	4.13	3.074	2.13	2.335	.37	.885	1.31	1.448

Figure 3. Effectiveness of WCF on subject-verb agreement over time

WCF was also successful in improving singular/plural errors, though a statistically significant difference is seen only for LOW and not for HIGH. Table 7 and Figure 4 show the data comparing LOW control to LOW experimental while Table 8 and Figure 5 show the data comparing both control groups to both experimental groups.

Table 7. Singular/plural errors (LOW control versus LOW experimental) (Descriptive statistics for mean test scores by group and elicitation task)

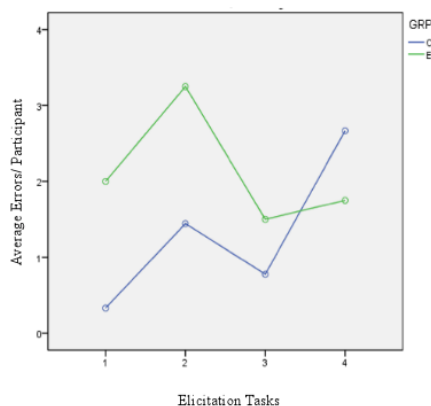| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	.33	.50	1.44	1.67	.78	.883	2.67	1.58
Experimentals	16	2.00	1.93	3.25	2.12	1.50	1.60	1.75	1.58

Figure 4. Effectiveness of WCF on singular/plural over time (LOW control versus LOW experimental)

Here we see that participants in the LOW control group noticeably increase their singular/plural errors over the four elicitation tasks, while participants in the LOW experimental group reduce their errors marginally. Still, the difference between the two LOW groups' behaviors regarding proper use of singulars and plurals across the elicitation tasks is significant (ANOVA F-Statistic=.05). The HIGH versus HIGH comparison was not statistically significant (ANOVA F-Statistic=.996). In fact, HIGH controls and HIGH experimentals were very similar in their abilities to reduce singular/plural errors over the 4 elicitation tasks.

Table 8. Singular/plural errors (LOW/HIGH controls versus LOW/HIGH experimentals) (Descriptive statistics for mean test scores by group and elicitation task)

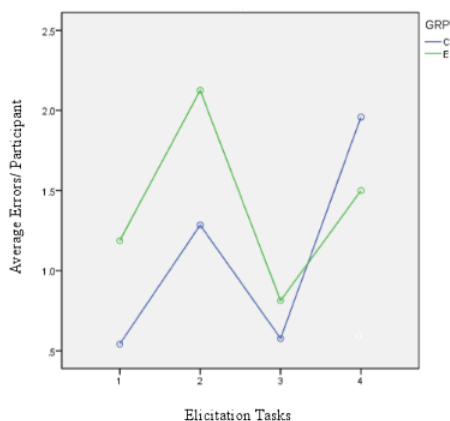| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	.53	.80	1.29	1.36	.59	.71	2.00	1.62
Experimentals	16	1.19	1.64	2.13	2.09	.81	1.33	1.50	1.37

Figure 5. Effectiveness of WCF on singular/plural over time (LOW/HIGH controls versus LOW/HIGH experimental)

Yet, when HIGH control and HIGH experimental data is combined with LOW control and LOW experimental data, singular/plural error behavior between the groups is quite similar. Nonetheless, results indicate that the correction symbol *num* can be beneficial in reducing singular/plural errors at lower proficiency levels. That is, WCF was instrumental in keeping LOW experimental participants from increasing singular plural errors like their LOW control counterparts.

4.3.2 Verb Tense and Articles

In contrast to the effectiveness of WCF in treating subject/verb agreement and singular/plural errors, we find no statistically significant differences between the groups for verb tense (ANOVA-F Statistic=.790) as seen in Table 9 and Figure 6 nor for articles (ANOVA F-Statistic=.127) as seen in Table 10 and Figure 7. That is, we obtain no positive data that WCF was effective on these error types, consistent with the notion that WCF was ineffective for these error types.

Table 9. Verb tense errors (Descriptive statistics for mean test scores by group and elicitation task)

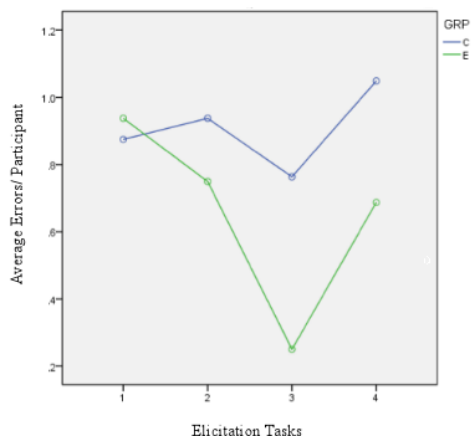| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	.88	1.11	.94	1.14	.76	1.20	1.06	1.68
Experimentals	16	.94	9.29	.75	1.52	.25	.45	.69	.95

Figure 6. Effectiveness of WCF on verb tense errors over time

We see here that experimentals are more successful at reducing verb tense errors from a percent reduction perspective (-27% between elicitation tasks 1 and 4) compared to controls (+20% between elicitation tasks 1 and 4). In fact, the controls in the study never outperform experimental groups in any of the individual error types from a percent reduction perspective. Of course, what really matters is statistical significance, which accounts for sample size and error count, and controls and experimentals do not perform differently enough to claim that verb tense is amenable to change using the *vt* correction symbol alone. The same appears to be true of the *art* symbol, as shown below.

Table 10. Article errors (Descriptive statistics for mean test scores by group and elicitation task)

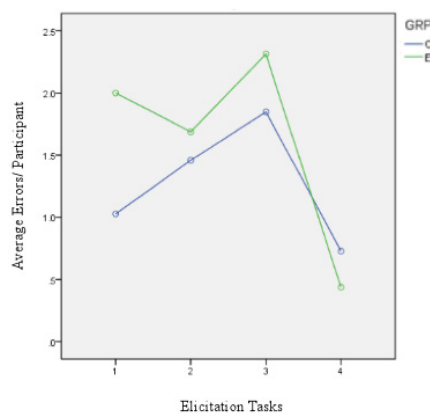| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Controls	17	1.06	.996	1.53	1.94	1.88	1.49	.76	1.20
Experimentals	16	2.00	1.50	1.69	2.63	2.31	1.92	.44	.89

Figure 7. Effectiveness of WCF on article errors over time

In Table 10 and Figure 7, we see that article error reduction trends between control and experimental groups are similar over the four elicitation tasks (ANOVA F-Statistic=.127). Thus, simply indicating the presence of article errors is not sufficient to guide students to acquiring the article system in English.

Thus, the answer to RQ3 is no. When a meaningful repair strategy can be supplied to a learner via WCF on a binary paradigm, WCF is effective. Otherwise, for non-binary paradigms, it is not.

4.4 Proficiency Levels

Results indicate that WCF functions similarly at both the LOW and HIGH proficiency levels (ANOVA F-Statistic=.364) as seen in Table 11 and Figure 8.

Table 11. Total errors of experimental groups by proficiency (Descriptive statistics for mean test scores by group and elicitation task)

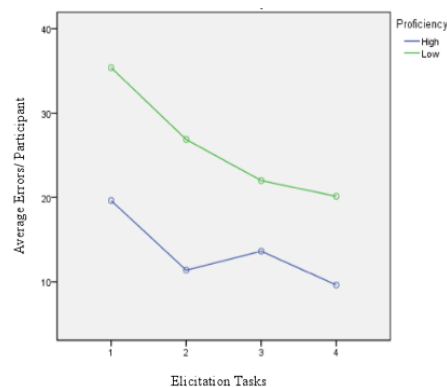| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
LOW Prof.	8	35.38	10.322	26.88	6.244	22.00	7.131	20.13	7.586
HIGH Prof.	8	19.63	12.258	11.38	5.528	13.63	8.088	9.63	5.854

Figure 8. Effectiveness of WCF on total errors over time by LOW and HIGH proficiency (experimentals only)

We see that while the LOW proficiency group makes more errors on average than the HIGH group (as expected) for each elicitation task, the two groups reduce errors over time comparably. The high F-statistic=.364, which is much greater than .05, means the groups are similar in error reduction. Consequently, it appears that indirect WCF is comparably effective overall for students at both proficiency levels.

Table 12 and Figure 9 compare the LOW and HIGH groups' abilities to reduce linguistic errors over time. Here again, both groups are similar in their abilities to use WCF to reduce errors over time (ANOVA F-Statistic=.150). Even though the LOW group has more errors for each elicitation task, both groups reduce errors similarly over the four tasks. This means that the type of indirect feedback used in this study can be beneficial to ELL students of both low and high proficiency in terms of overall error reduction.

Table 12. Linguistic errors of experimental groups by proficiency (Descriptive statistics for mean test scores by group and elicitation task)

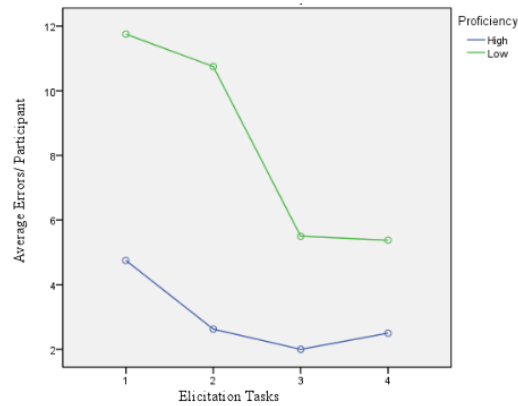| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
LOW Prof.	8	11.75	4.062	10.75	6.089	5.50	3.586	5.38	3.889
HIGH Prof.	8	4.75	3.845	2.63	2.200	2.00	2.000	2.50	2.204

Figure 9. Effectiveness of WCF on linguistic errors over time by LOW and HIGH proficiency (experimentals only)

WCF was similarly effective for reducing word-usage errors in both LOW and HIGH groups (ANOVA F Statistic=.533) as seen in Table 13 and Figure 10.

Table 13. Word-usage errors of experimental groups by proficiency (Descriptive statistics for mean test scores by group and elicitation task)

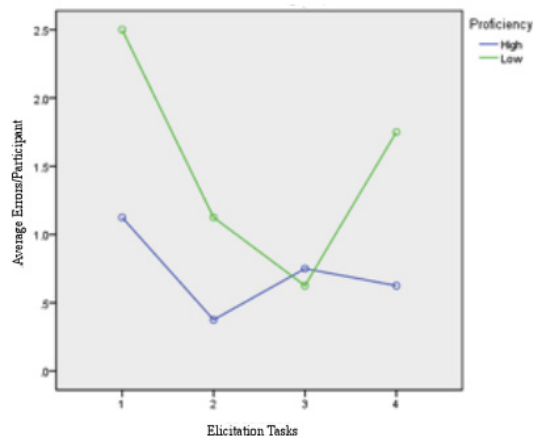| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
LOW Prof.	8	2.50	2.07	1.13	.99	.63	.74	1.75	1.04
HIGH Prof.	8	1.13	1.13	.38	.52	.75	.71	.63	.916

Figure 10. Effectiveness of WCF on word-usage errors over time by LOW and HIGH proficiency (experimentals only)

Finally, WCF was also similarly beneficial to both LOW and HIGH proficiency groups in reducing mechanical errors over the elicitation period (ANOVA F-Statistic=.749), as seen in Table 14 and Figure 11.

Table 14. Mechanical errors of experimental groups by proficiency (Descriptive statistics for mean test scores by group and elicitation task)

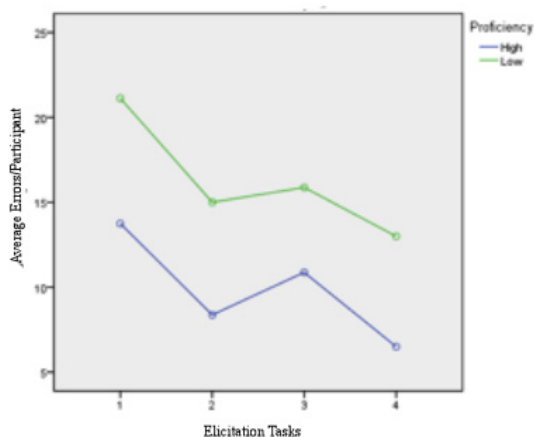| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
LOW Prof.	8	21.13	9.141	15.00	6.59	15.88	6.99	13.00	9.59
HIGH Prof.	8	13.75	9.30	8.38	5.21	10.88	7.62	6.50	5.37

Figure 11. Effectiveness of WCF on mechanical errors over time by LOW and HIGH proficiency (experimentals only)

These results demonstrate that the indirect WCF used here can be effective with both LOW and HIGH proficiency for overall errors, linguistic errors, word-usage errors, and mechanical errors. Consequently, the answer to RQ4 is yes.

5. Discussion

The results demonstrate that some individual error types are amenable to change via WCF whereas others are not, at least with the indirect WCF used here. The determining factors (which interrelate) are (1) having a binary versus non-binary paradigm and (2) the ability of a symbol to deliver an easy-to-learn repair strategy that can be utilized in future writing endeavors. When the right correction is not made explicit via WCF, the student might select another wrong form. This could impede learning a repair strategy for future writing endeavors. With a non-binary paradigm, correction symbols do not provide meaningful instruction or repair strategies in the form of positive information that students can learn. Under SAT, WCF is predicted only to work if it imparts sufficient declarative knowledge that can then be gradually transferred into procedural knowledge via practice. We can see how the results of the study are in line with the predictions of SAT about when WCF would be effective or ineffective.

Consider how students would interact with the subject-verb agreement symbol (*s-v*). Subject-verb agreement is indeed a relatively straightforward construction to repair (practice) and learn (proceduralize). The repair is easy: add this suffix or remove this suffix. Crucially, apart from some irregular forms, there are no other options to consider. Students know that the verb originally written must be incorrect, so the other option must be the correct one. Once the correction is made, the student knows the error is fixed. As a result, the practice of correcting the errors can help students transfer the declarative knowledge of subject-verb agreement to the proceduralized knowledge of using this pattern. Student interaction with the singular/plural symbol *num* is similar in that the repair is simply to convert an incorrect singular form to plural or vice versa.

By contrast, the symbol *vt* alone is not enough to provide meaningful instruction (i.e., imparting sufficient declarative knowledge) for the English tense system. It does not provide a repair strategy to employ in future writing. For verb tense, skill acquisition breaks down from the very beginning. Declarative knowledge is not imparted by the symbol, so practice is inconsequential and cannot facilitate a transfer from declarative to proceduralized knowledge. Students who receive the *vt* symbol know they are having verb tense problems, but

they can only guess at what the problem is. Even if they choose correctly from among the twelve verb tenses, they cannot be sure that they have made the right correction. If the student needs the present perfect tense, for example, it is not possible to extrapolate this from the symbol. Moreover, this student learns nothing about using the present perfect tense in order to express the tense's intended meaning.

Direct WCF (i.e., providing the correct revision) could possibly prove more effective for verb tense errors than indirect WCF given that the writing instructor can write which tense the student should have used and the reasons why that particular verb tense works better. Using direct feedback for verb tense errors should do precisely what Bitchener and Knoch (2010) review as some of the strengths of direct feedback, namely, "reducing confusion" and "resolving complex errors" (pp. 209-210). Indeed, it seems reasonable to assume that as linguistic complexity increases, the "directness" of WCF should also increase, though this hypothesis must be tested. If nothing else, the student is sure that the revision being made is correct, and this could enable the student to deduce at least some verb tense rules from the usage. The options within the English verb tense system are too many for indirect WCF to impart useful feedback via a single proofreading mark.

Similar to the *vt* symbol, when a student encounters the article symbol *art*, no instruction takes place, and no repair strategy is relayed to the student. A brief examination of the English article system demonstrates that we should not expect the symbol *art* to suffice. The symbol points out an article error, but offers no further clarification about how to repair the error. In the absence of sufficient instruction, students may draw overly simplistic conclusions about how the article system works. Indeed, indirect WCF instruction can be confusing to L2-English students who need to consider multiple rules concurrently in order to employ articles correctly.

Crucially, note that the article system involves a non-binary choice of options. To begin with, we have the definite article (e.g., *I bought the book*), but we have two forms for the indefinite article (e.g., *I bought a book*; *I bought an apple*). Beyond this, there are other choices in the paradigm. Another article option is the null article, meaning that no overt article whatsoever is used (e.g., *I bought books*). In addition, instead of using an article, it is possible to use a possessive form (e.g., *I bought my book*) or a demonstrative (e.g., *I bought this book*). We can also use a quantity expression with or without a definite article, but not with an indefinite article (e.g., *I bought the four books*; *I bought four books*; **I bought a four books*). Thus, although at first glance the article system may appear to be a binary choice between *the* versus *a/an*, there are indeed many more options involved.

Although it is not determined here what feedback will be useful in helping students learn the article system in English, a crucial first step is to recognize the ineffectiveness of attempting to teach this system via indirect WCF. Feedback on articles that provides the learner with sufficient declarative knowledge should be studied in more depth. In addition to straightforwardly supplying the correction with direct WCF, a system of mnemonics and diagnostics for teaching the article system in English, such as those argued for by Wulf (2016) could be tried. Indeed, mnemonic and diagnostic systems for other non-binary constructions besides the article system could be developed and evaluated.

6. Conclusion

This study supports a binary/non-binary hypothesis of grammar error WCF treatability instead of making claims about grammatical complexity. Establishing an objective measure of what constitutes a simple versus a complex grammatical construction in L2 learning has long been difficult to define. For example, Krashen (1982) claims English subject-verb agreement is simple based on the number of alternative forms (i.e., a binary choice), but Ellis (1990) regards it as complex based on the construction's processing demands of a long-distance relationship between the subject and the verb. DeKeyser (1998) regards the construction as complex, but bases this on its highly syncretic nature (i.e., it combines three abstract concepts—present tense, third person, singular). Thus, not only is there a lack of consensus on what constitutes simple and complex, but there are differences in how criteria associated with complexity have been applied in studies. However, by employing instead the notion of treatable and non-treatable errors, as proposed by Ferris (1999), it is possible to avoid the problem of fully conceptualizing grammatical complexity while still having clear criteria upon which to decide what kind of feedback is appropriate for addressing any given grammatical error. Namely, we must see if the error is in a grammatical paradigm that is binary or non-binary.

Our results also augment the plausibility of SAT and substantiate that a strong interaction can exist between explicit and implicit grammatical knowledge in SLA (Bialystok & Ryan, 1985; Sharwood-Smith, 1981; Anderson, 1983, 2005; Dienes & Perner, 1999; White & Ranta, 2002; Ellis R., 2009). That is, the results counter what is termed *the non-interface position* of SLA (Krashen, 1981; Krashen, 1982; Krashen & Terrell, 1983; Schwartz, 1993; Paradis, 1994, 2009; Hulstijn, 2002; Ellis R., 2004; 2009; Ellis N., 2005), which argues that no interaction can exist between explicit and implicit knowledge in SLA.

Because the binary pattern of treatable constructions has only two options to learn for correct usage, this allows declarative knowledge of these constructions to be induced from indirect WCF alone. With this binary option requirement, a WCF symbol not only points out and focuses students' attention on an error that has been made but also provides the correct answer by indicating the other option. Crucially, apart from perhaps some irregular forms in the paradigm, there are no other options within the construction for correct usage, even within varied semantic or pragmatic situations. Indeed, the binary composition of these treatable grammatical structures (e.g., including or omitting a suffix) coincides with the binary nature of WCF (i.e., correct or incorrect). Where there is no WCF, students assume there are no errors (i.e., correct). Where they receive WCF, an error has been made (i.e., incorrect). For binary constructions, providing the feedback is tantamount to providing the correct answer. In such cases, the other option is the correct one.

As we see, for a binary paradigm, WCF is akin to correcting errors flagged on a true/false test. Yet, by contrast, for grammatical constructions with more than a binary choice for correct usage, WCF is like correcting errors flagged on a multiple-choice test. That is, the feedback points out the error without providing any decipherable declarative knowledge for what the correction should be or why the correct correction is correct. In fact, there is no way for the L2 student to know if the right correction has been made in the subsequent draft.

Via SAT and the binary/non-binary hypothesis, it should be possible both to assess treatability cross-linguistically in an objective way as well as to predict for which constructions WCF will or will not be effective. Further examinations of this binary paradigm are warranted in order to examine whether it represents the most significant factor in determining WCF effectiveness.

In particular, it is plausibly not a strong enough condition since there are binary paradigms that have an idiosyncratic distribution. An example is Spanish masculine/feminine selection for individual words. Indeed, Wagner (2016) conducts similar studies of WCF effectiveness for both English and Spanish, and finds that for the idiosyncratic binary masculine/feminine paradigm in Spanish, indirect WCF is ineffective. This is in line with what Ferris (1999) assumes about all idiosyncratic paradigms. However, the present study and Wagner (2016) both demonstrate that not all rule-based paradigms are treatable via indirect WCF, as Ferris (1999) assumes. Thus, as we see, the (non-idiosyncratic) binary paradigm for WCF analysis provides us with a detailed understanding of a threshold of effectiveness for WCF. Similar criteria could also be used to examine other teaching strategies that practitioners assume impart sufficient declarative knowledge for skill acquisition to occur via practice.

Acknowledgements

References

Anderson, J. (1976). *Language, memory, and thought*. New York: Lawrence Erlbaum Associates.

Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. (2005). *Cognitive psychology and its implications*. New York, NY: Worth.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is the content feedback followed by form feedback the best method? *Journal of Second Language Writing*, *9*(3), 227-258. http://dx.doi.org/10.1016/S1060-3743(00)00027-8

Bartram, M., & Walton, R. (1991). *Correction*. Hove, England: Language Teaching Publications.

Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. In D. Forrest-Pressley, C. MacKinnon, & T. Waller (Eds.), *Metacognition, cognition, and human performance* (Vol. 1, pp. 207-252). Orlando, FL: Academic Press.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*, 102-118. http://dx.doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J., & Knoch, U. (2008). The value of written corrective feeback for migrant and international students. *Language Teaching Research*, *12*, 409-431. http://dx.doi.org/10.1177/1362168808089924

Bitchener, J., & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, *31*, 193-214. http://dx.doi.org/10.1093/applin/amp016

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, *19*, 207-217. http://dx.doi.org/10.1016/j.jslw.2010.10.002

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, *14*, 191-205. http://dx.doi.org/10.1016/j.jslw.2005.08.001

Bruton, A. (2010). Another reply to Truscott on error correction: Improved situated designs over statistics. *System*, *38*, 491-498. http://dx.doi.org/10.1016/j.system.2010.07.001

Chandler, J. (2003). The efficacy of various kinds of error feedack for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*, 267-296. http://dx.doi.org/10.1016/S1060-3743(03)00038-9

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

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 125-151). New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524780.007

DeKeyser, R. (2007). Introduction: Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 1-18). Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511667275.002

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735-808. http://dx.doi.org/10.1017/S0140525X99002186

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305-352. http://dx.doi.org/10.1017/s027226310505014x

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*(2), 227-275. http://dx.doi.org/10.1111/j.1467-9922.2004.00255.x

Ellis, R. (2009). *The study of second language acquisition*. Oxford, England: Oxford University Press.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, *36*, 353-371. http://dx.doi.org/10.1016/j.system.2008.02.001

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524551.016

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A reponse to Truscott (1996). *Journal of Second Language Writing*, *8*(1), 1-11. http://dx.doi.org/10.1016/S1060-3743(99)80110-6

Ferris, D. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...)? *Journal of Second Language Writing*, 49-62. http://dx.doi.org/10.1016/j.jslw.2004.04.005

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition*, *32*, 181-201. http://dx.doi.org/10.1017/S0272263109990490

Ferris, D., & Roberts, B. (2001). Error feedback in the L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, *10*(3), 161-184. http://dx.doi.org/10.1016/S1060-3743(01)00039-x

Ferris, D., Chaney, S. J., Komura, K., Roberts, B. J., & Mckee, S. (2000). Perspectives, problems, and practices in treating written error. In *International TESOL Convention*. Vancouver, B.C.

Guenette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, *16*, 40-53. http://dx.doi.org/10.1016/j.jslw.2007.01.001

Hendrickson, J. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal*, *62*, 216-221.

Higgs, T. (1979). Coping with composition. *Hispania*, *62*, 673-678. http://dx.doi.org/10.2307/340151

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193-223. http://dx.doi.org/10.1191/0267658302sr207oa

Hyland, K. (1990). Providing productive feedback. *ELT Journal*, *44*, 494-498. http://dx.doi.org/10.1093/elt/44.4.279

Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, *75*, 305-313. http://dx.doi.org/10.1111/j.1540-4781.1991.tb05359.x

Krashen, S. (1981). *Second language acquistion and second language learning*. London, England: Pergamon.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Englewood Cliffs, NJ: Prentice-Hall.

Krashen, S., & Terrell, T. (1983). *The natural approach*. New York, NY: Pergamon.

Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, *66*, 140-149. http://dx.doi.org/10.1111/j.1540-4781.1982.tb06973.x

McLaughlin, B. (1987). *Theories of second language acquistion*. London: Edward Arnold.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In *Implicit and explicit learning of languages* (pp. 393-419). San Diego, CA: Academic Press.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/sibil.40

Polio, C., Fleck, C., & Leder, N. (1998). If only I had more time: ESL learners changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, *7*, 43-68. http://dx.doi.org/10.1016/S1060-3743(98)90005-4

Raimes, A. (1983). *Techiniques in teaching writing*. Oxford, England: Oxford University Press.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, *20*, 83-93. http://dx.doi.org/10.2307/3586390

Schwartz, B. D. (1993). On explicit and negative evidence effecting and affecting competence and "linguisitic behavior". *Studies in Second Language Acquisition*, *15*, 147-163. http://dx.doi.org/10.1017/S0272263100011931

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, *17*, 195-202. http://dx.doi.org/10.1111/j.1944-9720.1984.tb01727.x

Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, *2*, 159-168. http://dx.doi.org/10.1093/applin/2.2.159

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, *41*, 255-283. http://dx.doi.org/10.1002/j.1545-7249.2007.tb00059.x

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, *32*, 203-234. http://dx.doi.org/10.1017/S0272263109990507

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, *23*, 103-110. http://dx.doi.org/10.1177/003368829202300107

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguisitic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, *22*(3), 286-306. http://dx.doi.org/10.1016/j.jslw.2013.03.011

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*, 327-369. http://dx.doi.org/10.1111/j.1467-1770.1996.tb01238.x

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*, 111-122. http://dx.doi.org/10.1016/S1060-3743(99)80124-6

Truscott, J. (2004). Evidence and conjecture: A response to Chandler. *Journal of Second Language Writing*, *13*, 337-343. http://dx.doi.org/10.1016/j.jslw.2004.05.002

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 255-272. http://dx.doi.org/10.1016/j.jslw.2007.06.003

Truscott, J., & Hsu, A. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*, 292-305. http://dx.doi.org/10.1016/j.jslw.2008.05.003

Wagner, J. (2016). *An explanation of the effectiveness of written corrective feedback in second-language grammar acquisition* (Ph.D. dissertation). George Mason University.

White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness*, *11*(4), 259-290. http://dx.doi.org/10.1080/09658410208667060

wida.us/standards/CAN_DOs/. (2009). *B. o. Consortium, Producer, & World-Class Instructional Design and Assessment*. Retrieved April 21, 2012, from http://www.wida.us

Wulf, D. (2016). Assistance in learning and teaching the English article system: Diagnostics and mnemonics. *Open Journal of Modern Lingustics*, *6*, 317-328. http://dx.doi.org/10.4236/ojml.2016.64032

Copyrights