2014-03-01

A Modified Approach to the Implementation of Dynamic Written Corrective Feedback

Brooke Elizabeth Eddington
Brigham Young University - Provo

A Modified Approach to the Implementation of Dynamic Written Corrective Feedback

Brooke Elizabeth Barton Eddington

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Norman W. Evans, Chair
Neil J Anderson
K. James Hartshorn

Department of Linguistics and English Language

Brigham Young University

March 2014

ABSTRACT

A Modified Approach to the Implementation of
Dynamic Written Corrective Feedback

Brooke Elizabeth Barton Eddington
Department of Linguistics and English Language, BYU
Master of Arts

Grammatical accuracy in second language (L2) writing is one of the key issues that English as a Second Language (ESL) learners struggle with, both in intensive English language programs and continuing after their university matriculation. Numerous instructional methodologies exist that center around the concept of error correction—how can or should ESL instructors correct grammatical errors in L2 students' writing to best facilitate improvements in written linguistic accuracy? Error correction in L2 writing has been a controversial issue for over a decade (e.g., Ferris, 1999; Truscott, 1996), and in an effort to contribute to an understanding of this controversial topic, this study investigated an innovative method of error correction known as dynamic written corrective feedback (WCF). For 15 weeks, 24 students at the Brigham Young University (BYU) English Language Center (ELC) received a form of dynamic WCF dramatically modified from Hartshorn's (2008) original method with the objective of increased practicality. These students produced a 30-minute pretest and posttest essay, and researchers calculated the complexity, accuracy, and fluency of each pretest and posttest. Data from the current study is compared against data from Hartshorn (2008), which found dynamic WCF to be successful in improving accuracy after carrying out similar research. The results validate previous findings and confirm that dynamic WCF is an effective approach to error correction, even when dramatically modified.

Keywords: ESL, L2, L2 writing, error correction, written corrective feedback, dynamic written corrective feedback, linguistic accuracy

ACKNOWLEDGEMENTS

This document represents the culmination of my graduate experience. What a truly fantastic experience it's been. The journey of writing this thesis would not have been possible without the support of many who have been with me throughout the process.

First and foremost, I express thanks to my Heavenly Father for blessing me with the opportunity to pursue a graduate education. I am grateful to my parents, Bruce and Stefani Barton, who have been the ultimate source of support in every possible regard. Thanks to my husband, Mike Eddington, for his continual confidence in me and his willingness to be a listening ear. I also express gratitude to my sister, Annie Barton, for the love and encouragement she has shown me in everything I've pursued in my life.

I am indebted to my chair, Dr. Evans, who inspired me as an undergraduate to pursue linguistics and ultimately TESOL. His vision and insight on this paper have been invaluable, as well as his mentorship and example to me throughout my university education. I am also immensely grateful to Dr. Anderson for always inspiring me and pushing me, and to Dr. Hartshorn for his enthusiasm about this research and statistical help.

Deserving of special mention is Judson Hart, a mentor without whom this research would not have been possible, as well as my friends and fellow colleagues in the TESOL MA for providing unfailing support. Ultimately, thank you to every influential person in my life who has inspired me and expressed confidence in me.

Table of Contents

List of Tables

List of Figures

Introduction

University-level learners of English as a second language (ESL) face a seemingly insurmountable task when it comes to writing: concurrent to mastering the rhetorical aspects of writing, they must also achieve grammatical accuracy in their production of written language. For over a decade, second language (L2) researchers have heatedly debated the effectiveness of error correction as a practice to improve L2 writing accuracy (e.g., Bruton, 2010; Ferris 1999, 2004; Truscott, 1996, 2007; Truscott & Hsu, 2008). Paramount in the discussion among researchers is whether error correction, commonly referred to as written corrective feedback (WCF), is valuable or detrimental.

Ferris (2004) observed that "the existing research base does not adequately address the big question: Does error feedback help L2 student writers?" (p. 50). This clear call for future research resulted in numerous studies investigating error correction. In light of significant evidence favoring WCF (Bitchener, 2008; Bitchener & Knoch, 2010; Bitchener, Young & Cameron, 2005; Chandler, 2003; Evans, Hartshorn, McCollum, & Wolfersberger, 2010; Evans, Hartshorn, & Strong-Krause, 2011; Ferris, 1999, 2004, 2006; Ferris, Liu, Sinha, & Senna, 2013; Guénette, 2007; Hartshorn, 2008; Hartshorn et al., 2010; Sheen, 2007), contemporary research seems to echo the sentiment of Evans et al. (2010) that "the WCF debate has been framed by the wrong question. . . . Rather than asking *whether* to provide WCF, the more essential question is *how* we help our students write more accurately" (p. 2).

In an attempt to clarify effective error correction, a group of researchers (Evans et al., 2011 & Hartshorn et al., 2010) created a systematic approach called dynamic written corrective feedback. Its instructional methodology adheres to four principles which ensure that writing tasks and feedback given are meaningful, timely, constant, and manageable. In the dynamic

WCF approach, students in an L2 writing classroom produce short, focused pieces of writing for 10 minutes daily. Immediately thereafter, the teacher provides indirect feedback on the piece of writing by marking coded symbols over productions containing errors. The following day, the marked draft is returned to the student by the teacher. The student corrects his or her own errors based on the indirect feedback provided, then rewrites the draft again. This process continues until the student produces a draft that is error-free. In an L2 writing classroom implementing dynamic WCF, daily grammar instruction focuses on specific errors produced by the students.

In a treatment-control group research design, Hartshorn (2008) found dynamic WCF to be a successful approach to error correction. However, Brown and Larson-Hall (2012) claim that "there are no new studies that show that correcting all of the errors in a piece of student writing is beneficial . . . as such studies have been rare" (p. 116). This statement ignores the recent—although limited—literature on dynamic WCF, of which there is no evidence contradicting its effectiveness (Evans et al., 2010, 2011; Hartshorn, 2008; Hartshorn et al., 2010; Hartshorn & Evans, 2013). Thus one purpose of the current study is to contribute further research supporting dynamic WCF as an effective approach to error correction, and eradicate claims such as Brown and Larson-Hall's that new studies on error correction are rare.

This study also seeks to contribute insight into researchers' present understanding of that quintessential *how* question of error correction (Evans et al., 2010). Considering the success of dynamic WCF, this study investigates the effect of a dynamic WCF instructional methodology dramatically modified from how it is presented in Hartshorn (2008) on L2 students' writing accuracy. The modified instructional methodology primarily seeks to make the dynamic WCF method more practical and approachable for practitioners without negating students' improvements in accuracy. The results of Hartshorn (2008) and the current study are compared.

Literature Review

Both a description of WCF and a brief review of its presence in the literature are necessary to contextualize this study. Following is an exploration of various WCF methods, the ensuing controversy surrounding error correction, and reasons for contradicting views of its effectiveness in the literature. A discussion of dynamic WCF is also presented, as well as the research questions for this study.

Written Corrective Feedback

At its core, WCF is error correction given by a teacher with the objective to improve a student's L2 writing accuracy. Since there are numerous variations of error correction, conclusions made in the literature about its effectiveness are diverse and thus many researchers and teachers view WCF as a controversial issue. Whether WCF has adverse or positive consequences on writing accuracy is central to the debate. Indeed, "responding effectively to students' grammatical and lexical problems is a challenging endeavor fraught with uncertainty about its . . . effectiveness" (Ferris, 1999, p. 1).

There are four basic approaches to WCF, which have been developed and researched in attempts to either discover an effective error correction method or give rationale for abandoning it altogether. Broadly speaking, the variations in the way WCF is implemented can be summarized as direct, indirect, focused, or unfocused.

Direct and indirect feedback. Hartshorn (2008) references Ferris's (2006) definitions of these two different types of WCF that can be given on L2 writing: "direct feedback is provided when a teacher gives students a particular correction and indirect feedback is provided when the teacher simply marks the error but does not correct it" (p. 29). Many "studies have compared direct and indirect feedback methods. . . [but] there is no common conclusion about the findings.

Some . . . indicate that direct feedback is more effective whereas other studies emphasize the effectiveness of indirect feedback over direct feedback" (Farid & Samad, 2012, p. 234).

Ferris and Roberts (2001) carried out a study comparing two treatment groups receiving some form of indirect WCF with a control group receiving no WCF. The first treatment group had errors underlined and coded, the second treatment group had errors underlined with no coding, and the control group received no markings whatsoever. From the results, it is "evident that both groups which received error feedback substantially outperformed the control . . . group" (p. 171). Between the treatment and control groups, "the differences in editing success are quite striking. In contrast, there were no statistically significant differences in editing ratios between the "codes" and "no codes" groups" (p. 172). In addition to the overall effectiveness of indirect WCF, the researchers suggested that "indirect feedback can . . . help students to self-edit idiosyncratic errors such as word choice and sentence structure" (p. 172). Direct feedback does not promote self-awareness, while indirect feedback does.

According to Ferris and Roberts (2001), indirect feedback is generally preferable for students because it is cognitively engaging as students are directly involved in problem solving and self-guided learning. Furthermore, indirect feedback forces students to become aware of their own errors and to self-correct because they are not supplied with the answer. When students do not expend any effort to self-correct, as in the direct feedback model, this will not foster long-term acquisition. Ferris and Roberts summarize research on direct and indirect error correction by stating that "indirect feedback helps students to make progress in accuracy over time more than direct feedback does" (p. 164). Ferris et al. (2013) posit that "indirect WCF may be more valuable for the long term than direct" (p. 309).

The instructor in the current study used a marking method of coded, indirect feedback on the student writing produced for this study. As pointed out by Ferris and Roberts (2001), choosing to use indirect WCF raises the question of "how explicit indirect feedback should be in order to give students enough direction to self-correct their errors" (p. 164). Ferris and Roberts continue to discuss that there are two basic options: highly explicit error correction, which marks the error at its exact location and labels it with a code, and more vague correction, such as placing a checkmark in the margin to signify an error somewhere on that line, but not specifying exactly where on the line. The current study used highly explicit WCF, marking errors directly over the text and labeling them with a code.

Focused and unfocused WCF. According to Sheen, Wright, and Moldawa (2009), focused WCF is "targeting only one linguistic feature" (p. 559). Unfocused WCF, then, is the opposite: error correction that targets all linguistic features equally.

Some research has focused exclusively on the accuracy of specific grammatical elements of L2 students' writing marked with WCF, for example, articles. Among three treatment groups and one control group (n = 75) in Bitchener's (2008) New Zealand study, the three treatment groups receiving focused WCF on articles showed drastic improvements in their article accuracy when compared to the control group. Bitchener and Knoch (2010) conducted similar research in the U.S. using the same research design with three treatment groups and one control group (n = 63). Again, the three treatment groups receiving focused WCF on articles outperformed the control group in article accuracy. Sheen (2007) also provided focused WCF on articles, drawing the same conclusion as Bitchener (2008) and Bitchener and Knoch (2010): from pretest to posttest, the two treatment groups receiving error correction showed marked gains in article accuracy when compared to one control group receiving no form of error correction treatment.

While focused feedback is indeed effective, by its nature it cannot address the range of grammatical errors students make in one piece of writing and thus is effective only within its narrow scope. This leads naturally to an examination of unfocused feedback and its efficacy in comparison with that of focused feedback. Some researchers criticize unfocused feedback, stating that it is "cumbersome, for both teachers and writers, and more complicated" (Bruton, 2010, p. 495). Sheen et al. (2009) reject unfocused feedback, stating that "when the correction addresses a range of grammatical errors, learners are unable to process the feedback effectively" (p. 565). Bruton summarizes the argument of Ellis, Sheen, Murakami, and Takashima (2008) that "maybe it is more realistic to consider degrees of focus, in terms of the number of features focused on" (p. 495) rather than a strict, polarized dichotomy of focused or unfocused.

However, a body of research (Evans et al., 2010, 2011; Hartshorn, 2008; Hartshorn et al., 2010; Hartshorn & Evans, 2013) has utilized WCF that leans towards the unfocused side of the spectrum with positive findings. Lee (2009) recommends using unfocused feedback, stating that teachers generally prefer focused feedback because it reduces their workload, but "treating some categories of errors does not satisfy students' actual needs and can be considered impractical" (p. 20). The current study also used unfocused WCF: in other words, it considered a wide range of errors and marked them without focusing on one specific error category.

It is important to note that the feedback given in the current study was not entirely unfocused according to definitions in the literature such as Ferris et al. (2013): "comprehensive . . . correction of *all* errors noticed by the teacher or researcher" (p. 309, emphasis added). To this end, the feedback given in this study falls somewhere along the continuum suggested by Ellis et al. (2008)—while it certainly falls on the unfocused side of the continuum, it is not

entirely removed from a certain degree of focus as the instructor in this study marked errors according to an extensive, though not comprehensive, list of predetermined error categories.

Contradicting Views of WCF

As different researchers employ the different methods of error correction previously discussed, various and contradicting results emerge from studies that lend to a confusing and unclear view of whether or not WCF is actually effective at improving linguistic accuracy. Following is a brief review of studies and researchers that believe WCF to be harmful, as well as those that claim WCF is helpful.

WCF is harmful. Some studies (Truscott, 1996, 1999, 2007; Truscott & Hsu, 2008) conclude that WCF is not beneficial to students and state various reasons as to why this is the case. Truscott believes WCF to be a "clear and dramatic failure" (2007, p. 271), claiming it is futile; ineffective; even harmful—"grammar correction has no place in writing courses and should be abandoned" (1996, p. 328). He asserts that the question "How effective is correction?" should be replaced by "How harmful is correction?" (2007, p. 271). Regarding the reportedly harmful nature of error correction, Truscott declares that "learning is most successful when it involves only a limited amount of stress, when students are relaxed and confident and enjoying their learning; but the use of correction encourages exactly the opposite condition" (1996, p. 354).

In addition to suggesting that WCF is harmful, Truscott (1996) claims there are other negative consequences of WCF. While discussing limitations in previous research that may have led to a non-improvement scenario in WCF studies and impeded the effectiveness of WCF, Truscott notes that error correction generally ignores the order, or sequence, in which grammatical structures are naturally acquired in L2 learning. A grammar principle is not

mastered suddenly or immediately, but rather mastery is the result of a patient and gradual building process. From this perspective, students cannot learn to be accurate if WCF comes developmentally prior to their mastery of one grammatical concept or another. As such, the "negative results in [prior] studies could have been due not to problems inherent in correction but rather bad timing" (p. 336). There are "clear and consistent orders in which learners acquire certain grammatical structures" and when "instructional sequences run counter to them . . . this raises the possibility that the corrections used in [research] failed because they did not respect these sequences" (pp. 336-337). Certain grammar structures may need to pedagogically precede what is being corrected in WCF. There is a sequence which must be followed, and students will not understand WCF at an inappropriate time in their acquisition process. Throughout much of the research he analyzed, Truscott noticed that "teachers corrected students on grammar points for which they were not yet ready" (p. 337).

Truscott (2007) criticizes studies that determined WCF was helpful in improving L2 writing accuracy, claiming that researchers have ignored the possible effect of avoidance strategies. Such avoidance strategies may be employed by students fearful of receiving feedback, and according to Truscott, this is a shortcoming and negative outcome of WCF. It is possible that improvements in L2 writing accuracy as perceived by researchers do not exist in reality, but that "corrected students hide their weaknesses. . . .When their scores rise on overall accuracy, this apparent improvement might simply mean they have learned to avoid using things they might get wrong" (p. 268). Because students dread feedback and fear making mistakes, they will embrace structures they are confident in and experience success with, and avoid structures that are difficult for them or are frequently corrected by the instructor. This results in WCF having little or no benefit and produces students who don't experiment or take risks with language. Students

are conditioned to be reluctant to attempt complex writing because they are fearful of feedback. WCF causes students to write in simple tenses and produce simple forms below their potential, and the only reason teachers continue to use WCF, according to Truscott, is because their intuition says it must be useful and that something is amiss with current research.

WCF is helpful. On the opposite side of the debate are other researchers who argue in favor of WCF and its positive outcomes in improving L2 writing accuracy. An abundance of studies exist that proclaim benefits of WCF and affirm its necessity in a writing accuracy classroom (Bitchener, 2008; Bitchener & Knoch, 2010; Bitchener, Young & Cameron, 2005; Bruton, 2010; Chandler, 2003; Evans et al., 2010; Evans et al., 2011; Ferris 1999, 2004, 2006; Ferris et al. 2013; Hartshorn, 2008; Hartshorn et al., 2010; Sheen, 2007). These researchers argue on the contrary to claims such as Truscott's, criticizing inconsistencies in research design and noting limitations in studies where WCF was deemed a failure.

Ferris (1999) writes a direct rebuttal to Truscott (1996), in which she criticizes his definition of WCF as well as his lack of supporting evidence. Truscott stated that "correction comes in many different forms, but…such distinctions have little significance" (p. 329), a definition with which Ferris disagrees. Ferris asserts that

> as with any other aspect of teaching, there are more and less effective ways to approach error correction in L2 writing. We would all doubtless agree that poorly done error correction will not help student writers and may even mislead them. However, there is mounting research evidence that effective error correction—that which is selective, prioritized, and clear—can and does help. (p. 4)

This suggests that there are ways that WCF can be done effectively, and distinctions among the implementation of, instructional methodology behind, or rationale for feedback given are not

trivial or irrelevant. As the current study strives to further clarify, distinctions between WCF instructional methodologies can make significant differences in student progress.

Polio (2012) delves more deeply into Truscott's (1996) claims that WCF ignores the natural order of acquisition in L2 learning. She identifies six theories of second language acquisition (SLA): usage-based, skill acquisition, sociocultural, interactive, generative, and processability. Her research investigates and discusses each theory in depth and its relation to WCF. Polio suggests that "error correction is not completely useless or harmful" (p. 384), and that the majority of SLA theories (usage-based, skill acquisition, sociocultural, and interactive) support its role in improving L2 writing, while only the two theories of generative and processability do not.

Ferris (2004) claims that most research into WCF has "overlooked or understated some potentially positive research evidence on the effects of grammar correction" (p. 50) and that relatively few studies have been done comparing a group receiving WCF to a group from which WCF was being withheld. She believes the lack of such studies is due to teachers feeling it is unethical to withhold WCF if their intuition tells them it is a beneficial practice, so they only do research on groups receiving WCF and compare the end result to the starting point. Without a control group, the comparison of gains in accuracy might seem insignificant between a pre- and posttest. Ferris argues that with a control group in a correction/no correction research scenario, accuracy improvements are apparent and significant.

Chandler (2004) "certainly agree[s] that evidence on the efficacy of error correction comes from comparisons between the writing of students who have done error correction and those who have not," but she acknowledges alongside Ferris (2004) that "it is difficult for teachers to give no feedback because of the strong desire of most students to have corrective

feedback" (p. 345). Guénette (2007) further confirms that researchers "must compare students who have received grammar correction with students who have not" (p. 43), despite feelings that withholding correction is unethical. Having a control group and a treatment group for comparison would "confirm that the time spent correcting the students' errors is not in vain" (p. 44).

In addition, Ferris (2004) states that error correction cannot be deemed as effective or ineffective because all research addressing it is "fundamentally incomparable because of inconsistencies in design" (p. 50). Guénette (2007) agrees, suggesting that an examination of the research designs and extraneous variables affecting the outcomes of all WCF studies is necessary to understand why these studies have produced such disparate results. The first weakness Guénette identifies in making comparisons across WCF studies is that generally this means comparing groups of different proficiency levels: "steps were usually taken to ensure that no significant differences existed between groups in proficiency level, [but] these steps were . . . not as rigorous as they needed to be" (p. 42). The "overall proficiency level of the students must be considered" and since the students' proficiency levels in most studies "were either not carefully measured or reported, it is nearly impossible for other researchers to replicate the study" (p. 43), make comparisons, or draw accurate conclusions about WCF effectiveness or ineffectiveness.

Guénette (2007) states that "the treatment is the crux of the matter" (p. 45). Current research studies have administered a wide range of treatments; thus, the results of these studies cannot be justly compared. Differences exist in the types of WCF feedback provided across studies: form only vs. form and content, direct vs. indirect, or focused vs. unfocused. The way that data is elicited is inconsistent, as well as the frequency of data elicitation and the amount of feedback students receive—every day, once every two weeks, or three times per semester will

produce drastically different outcomes. Guénette poses the question: "If students . . . are engaged in different classroom activities . . . how can the effects of feedback be isolated?" (p. 49).

Many studies differ in their research design in that they are either longitudinal or cross-sectional, and thus cannot rightly be compared. Additionally, a "major weakness of many experimental studies" is that "there is no report of interrater reliability" (Guénette, 2007, p. 42), or that one researcher analyzed all the data independently. Furthermore, no studies offered the same student incentives, which may have affected the results because students were receiving points or a grade based on different measures (e.g., form, content, or word count).

Ultimately, most studies on WCF are "not necessarily comparable because the design and methodology were not constant . . . [these] are indeed at the root of the different results obtained" (Guénette, 2007, p. 51). Therefore, the "lack of positive effects for written corrective feedback shown in many studies might be seen as inconsistencies in the research design, rather than as evidence that feedback does not work" (p. 41).

In a recent study, Ferris et al. (2013) found that students who are themselves the recipients of WCF perceive it as being highly effective. She conducted extensive interviews with students, and concluded that they consider WCF to be relevant, clear, and motivating. After gathering extensive anecdotal evidence of why WCF is effective from students receiving the treatment, Ferris explains that "students found the individualized and interactive teaching and learning provided through the feedback . . . to be extremely valuable" (p. 322). She urges teachers to "consider how they might fine-tune their own feedback processes" (p. 322). If these suggestions are followed, Ferris proposes that WCF will be effective in improving L2 writing accuracy.

Inconclusive findings. Consider the following two research findings as underpinnings of the controversy and reasons behind the development of various WCF instructional methodologies. First, Truscott and Hsu (2008) found that WCF improved accuracy only on immediate revisions of the same draft which received the markings. One week after the revisions, however, students could not remember their mistakes and repeated them again. In other words, students who received WCF performed just as poorly on a new writing task as did students who had received no WCF. Second, Bitchener (2008) reported that L2 writing students receiving WCF for a semester outperformed those in a control group not receiving WCF on accuracy. Contrastingly to Truscott and Hsu, the same held true on a delayed posttest administered two months after the end of the semester, implying that students indeed retained a memory of their errors and were able to perceive or self-correct them two months later.

Such conflicting findings perplex writing instructors and create an unclear view of WCF and its place in an institution. Clearly, this is an important field of research as many are seeking clarification and an error correction instructional methodology that is effective. In the words of Ferris (1999), there is an "urgent need for new research efforts which utilize a variety of paradigms to examine a range of questions that arise around this important topic" (p. 2).

Dynamic Written Corrective Feedback

In response to the controversy and in an effort to refine WCF and make it maximally effective, a group of researchers developed dynamic written corrective feedback (Evans et al., 2011; Hartshorn, 2008; Hartshorn et al., 2010). Dynamic WCF is defined by Hartshorn et al. (2010) as having

> two essential elements . . . (a) feedback that reflects what the individual learner needs most, as demonstrated by what the learner produces, and (b) a principled approach to

> pedagogy that ensures that writing tasks and feedback are *meaningful*, *timely*, *constant*, and *manageable* for both student and teacher. (p. 87)

Evans et al. (2011) clarify that the word *dynamic* was deliberately chosen for inclusion in naming this instructional methodology because of the interactive and continuous nature of feedback as it adheres to the four principles.

In the dynamic WCF approach, students in an L2 writing classroom produce short, focused pieces of writing for 10 minutes daily by responding to a prompt provided by the teacher. Immediately thereafter, the teacher provides indirect feedback on the piece of writing by marking coded symbols over productions containing errors. The following day, the marked draft is returned to the student by the teacher. The student corrects his or her own errors based on the indirect feedback provided, then rewrites the draft again. This process continues until the student produces a draft that is entirely accurate. It may take numerous rewrites for a student's draft to be completely error-free, but students are limited to however many rewritten drafts they can produce in a one-week period. In an L2 writing classroom implementing dynamic WCF, daily grammar instruction focuses on specific errors produced by the students.

Brown and Larson-Hall (2012) state that "haphazardly correcting errors…is not an effective way for a teacher to spend time with their students" and that "there is a place for effective correction, but it has to be quite conscious and deliberate and sustained over a period of time" (p. 107). Dynamic WCF fulfills these criteria and answers the call for a deliberate approach that is sustained over a period of time. Additionally, Brown and Larson-Hall write that "some correction [is] useful, especially if it [is] . . . given consistently" (p. 115). This fits the dynamic WCF principle of constant: feedback is consistently provided on a daily basis.

Hartshorn (2008, et al. 2010) carried out a study on dynamic WCF which is closely modeled in the current study. Dynamic WCF was analyzed, as implemented at the Brigham Young University (BYU) English Language Center (ELC) in Provo, Utah, United States. Forty-seven students participated: 28 were in a treatment group receiving dynamic WCF; 19 were in a control group receiving traditional process writing instruction. The objective was to determine whether dynamic WCF produced significant gains in accuracy among the treatment group students when compared to students in the control group receiving instruction unfocused on accuracy. Both groups received instruction for a 15-week semester, and teachers closely followed the instructional methodology prescribed for their group. The treatment group received tailored instruction based on diagnostic information of their specific grammar needs. The control group received formal instruction on writing conventions and general process writing, void of consideration to their individual needs (e.g., paragraph development; transitions; essay structure).

Approximately four days a week for 15 weeks, students in the treatment group wrote 10-minute compositions that were indirectly marked for accuracy using specific symbols (Appendix A). No feedback on rhetorical aspects of writing was provided, but students received a score on each composition out of 10, weighted 75% for accuracy and 25% for content. Students in the control group wrote four major papers throughout the 15-week course, which included multiple drafts that were marked for both rhetorical functions and accuracy.

Researchers administered a pretest 30-minute essay to both groups at the beginning of the semester prior to instruction, and a posttest 30-minute essay to both groups at the end of the semester after instruction. Three raters rated the 47 pretest and posttest essays (total = 94) in a fully-crossed design, examining accuracy, fluency, complexity, and rhetorical competence. The researchers were specifically attentive to measuring gains in accuracy and possible decreases in

fluency, complexity, or rhetorical competence, taking into consideration Skehan's (1998) notion that trade-offs might potentially occur: a gain in one area may compromise another, resulting in regression in some areas in spite of progression in others.

Ultimately, the treatment group outperformed the control group in accuracy by showing a significant increase (p = .001, η_p^2 = .21) from pretest to posttest, while the control group showed a decrease in accuracy. Rhetorical competence was virtually unaffected, with slight non-significant improvements in both groups (p = .77, η_p^2 = .002). Fluency improved somewhat more for the control group, but it also improved for the treatment group with negligible significance (p = .19, η_p^2 = .04). Writing complexity decreased in the treatment group but increased in the control group (p = .079, η_p^2 = .067), aligning with the sentiments of Skehan (1998) that increases in one area (accuracy) may lead to decreases in another (complexity) (as cited in Hartshorn et al., 2010). The results of Hartshorn's (2008) study show significant improvement in L2 writing accuracy, suggesting the dynamic WCF instructional methodology was highly effective regarding accuracy.

Research suggests that there is a key constraint to the successful implementation of dynamic WCF. Lee (2009), who abbreviates dynamic WCF even further to DWCF, "attempted to examine the effectiveness of DWCF in terms of the proficiency level variable" (p. 62) and advises that dynamic WCF is effective at advanced levels but perhaps not effective in intermediate and lower levels. Lee found that "more proficient students can benefit more from DWCF" (p. 64) and that at intermediate levels, "students improved their writing accuracy when taught either with [DWCF] or with a traditional grammar instruction method" (p. 65). At advanced levels, however, students in a control group receiving traditional grammar instruction showed decreases in overall linguistic accuracy while a treatment group receiving dynamic WCF

showed drastic and statistically significant increases. Hart (2011) agrees with Lee that advanced learners will likely increase linguistic accuracy as a result of dynamic WCF. From his study on dynamic WCF, Hart concludes that in an intensive English-learning context with high institutional support, gains in accuracy will predictably come if teachers follow principled error correction instructional approaches. Dynamic WCF fits the description of a principled approach and facilitated improvements in accuracy of advanced L2 student writing in Hart's study.

Evans et al. (2011) investigated improvements in accuracy over the course of one semester among a control and treatment group of university-matriculated ESL students at BYU. The results suggest that dynamic WCF had a positive effect on the writing accuracy of the treatment group with negligible decreases in fluency and complexity, while the control group unexposed to dynamic WCF experienced a decrease in their accuracy over the semester. This study demonstrates that dynamic WCF is effective in multiple contexts: not only in intensive English language-learning programs, but also after ESL students' matriculation into a U.S. university.

Of course, as with virtually any approach to error correction in the literature, the dynamic WCF method is not without its critics. McQuillan (2012) reviewed Evans et al. (2010), Evans et al. (2011), and Hartshorn et al. (2010). After an extensive analysis of all three studies, McQuillan concludes: "Unlike the researchers, I do not find the practical effect of this huge investment of time in error correction very impressive." The author states that dynamic WCF requires "painstaking and massive error correction efforts," and that the method lacks practicality because only massive improvements in accuracy could make it truly worthwhile. Although they were statistically significant, the author believes that the increases in linguistic accuracy in these three studies were not great enough to justify the amount of time and effort expended. McQuillan

finishes his argument by describing how the amount of time invested in the dynamic WCF process "appears to be substantial" for both students and teachers, and "to leave the semester still making so many errors can hardly be claimed as a victory for error correction."

The current study. Despite the compelling effectiveness of dynamic WCF across the literature, arguments such as those posed by McQuillan (2012) are valid to a certain extent in the eyes of current practitioners in the same intensive English language program where Hartshorn's (2008) study took place. The concern of practicality and manageability has fueled instructors with the objective of increasing their students' written grammatical accuracy to experiment with variations of dynamic WCF that are less time-intensive.

The current study looks at one such variation of dynamic WCF that is implemented by instructors in this intensive English language program today. While these instructors still maintain the core of dynamic WCF as an "interactive strategy [that] adhere[s] to four principles to ensure that the feedback is meaningful, timely, constant, and manageable" (Evans et al., 2011, p. 232), they strive to reach the same end objective of increased linguistic accuracy in ways dramatically modified from dynamic WCF as it was originally conceptualized by Hartshorn (2008, et al. 2010), hearkening back to Ferris's (1999) argument against Truscott (1996) that distinctions between methodologies do matter greatly.

As such, the current study examines one dynamic WCF instructional methodology dramatically modified from how it was implemented in Hartshorn (2008, et al. 2010) and its efficacy in improving L2 writing accuracy. The instructor implementing this method of dynamic WCF will be referred to by the pseudonym "McKay" throughout the remainder of this text. The current study strives to replicate Hartshorn (2008) as closely as possible, using much of the same research design in the same context; however, McKay's modified dynamic WCF approach is

implemented in the classroom. McKay's dramatic modifications affect a variety of aspects of dynamic WCF, but the primary purpose of all modifications, according to McKay, was to improve the dynamic WCF elements of practicality and manageability from a practitioner's perspective.

Dynamic WCF as defined by Hartshorn (2008) is modified by McKay in the following notable areas: recycling and manageability of writing for both teacher and student; the prompts students respond to; the building of prompt context and establishment of background knowledge; the coding symbols used to mark grammar errors in writing; non-dichotomous classroom instruction focused both on rhetorical functions of writing and editing grammar mistakes; and the out-of-class learning and assignments students were engaged in. The intended paramount focus of all modifications is manageability: making the dynamic WCF process more practical approachable for L2 writing accuracy instructors. Each of these modifications will be discussed in greater detail later in this work.

Research Questions

Recall back to the *how* question of error correction, which essentially concerns how WCF should be practiced in a writing accuracy classroom to be maximally effective. Evans et al. (2010) assert that "until research answers this essential 'how' question, many teachers may continue to feel confused as they struggle to identify best practices for their specific classroom contexts" (p. 2). This dilemma guides the research questions for the current study, which compares Hartshorn's (2008) and McKay's instructional methodologies, ultimately seeking to contribute to our understanding of dynamic WCF as a method for improving linguistic accuracy and how it can be effectively implemented. The research questions that follow are adapted from Hartshorn (2008) and Hartshorn et al. (2010):

1. Based on 30-min pretest and posttest essays, will mean accuracy scores in a group experiencing a dynamic WCF treatment dramatically modified from Hartshorn (2008) increase, decrease, or remain the same as scores obtained by students in the Hartshorn study?
2. Based on 30-min pretest and posttest essays, will there be significant compromises in fluency or complexity?

Methodology

This section will articulate the research methods of the current study, including how this study was designed and carried out to answer the research questioned outlined previously. The context of and participants in the study will be described, as well as all experimental interventions and manipulations. A principal purpose of this section is to provide a clear understanding of how McKay's instructional methodology differed from that of Hartshorn's (2008) in relation to the principles of dynamic WCF. Finally, this section will describe data elicitation and analysis procedures.

Participating Program, Students, and Teacher

This study took place in the same intensive ESL program as Hartshorn's (2008) study. Students who attend this program are enrolled in four 65-minute classes per 15-week semester: writing, reading, listening/speaking, and linguistic accuracy (applied grammar). While the students in the current study were enrolled in the standard four classes, research took place solely within the linguistic accuracy class, as that is where dynamic WCF is employed.

Within this intensive English language program, students are placed in one of eight levels, ranging from a beginning-low Foundations Prep (level 1) to an advanced-high Academic C (level 8). Classes are divided into two proficiency levels: Foundations (levels 1-4, referred to as Prep, A, B and C) and Academic (levels 5-8, referred to as Prep, A, B, and C). These levels correspond roughly to proficiency levels established by the American Council on the Teaching of Foreign Languages (ACTFL) (American Council on the Teaching of Foreign Languages, 2012). Levels 1—4 (Foundations Prep—Foundations C) correspond with ACTFL proficiency levels of Novice Low (NL)—Intermediate High (IH), and levels 5—8 (Academic Prep—Academic C) correspond with Intermediate High (IH)—Advanced High (AH).

Twenty-four students participated in the current study: 22 enrolled in Academic B (level 7) at an ACTFL (2012) Advanced Low (AL)/Advanced Mid (AM) proficiency level, and two enrolled in Academic C (level 8) at an AM/AH level. Students were placed in these levels based on their score on an institutional level placement test. The students at these levels were chosen for this study to parallel the 28 treatment students in Hartshorn's (2008) study, who were of equivalently advanced proficiency status. As posited by Lee (2009), dynamic WCF is only effective with advanced students.

The 24 students in this study consisted of 14 females and 10 males, ranging from age 17-34 (mean 22.79). The students came from 12 countries (Brazil, Chile, China, Colombia, Ecuador, Guatemala, Hungary, Korea, Mexico, Peru, Taiwan, and Venezuela) and spoke the following native languages: Chinese (2), Hungarian (1), Korean (4), Portuguese (1), and Spanish (16). Compare these demographics to Hartshorn's (2008) treatment group (n = 28), with 16 males, 12 females, and speakers of the following languages: French (1), Japanese (2), Korean (6), and Spanish (19).

One primary instructor, McKay, taught all 24 students participating in the current study. The instructor was very familiar with dynamic WCF and its principles, and had five years of teaching experience. In the current study, McKay implemented dynamic WCF in his advanced-level linguistic accuracy classroom, but with dramatic modifications to the instructional methodology practiced by teachers in Hartshorn's study.

Essentially, McKay had been implementing dynamic WCF in numerous advanced-level linguistic accuracy classes for several years and had observed how the process naturally transformed from theory into practice for him and other instructors. In his observation, dynamic WCF (as outlined and presented in Evans et al., 2010, 2011; Hartshorn, 2008; Hartshorn et al.,

2010) proved to be practically difficult for him and other instructors teaching a linguistic accuracy class. Based on his own feedback about the process and feedback from other instructors, he perceived the need for potential modifications to make the dynamic WCF model less overwhelming and more practical for instructors. He therefore modified the original system dramatically and developed a significantly different instructional methodology. His modifications, as next described, differ noticeably from how dynamic WCF was implemented originally in Hartshorn (2008).

Experimental Manipulations and Interventions: Critical Differences between two Instructional Methodologies

This section will describe how McKay drastically modified the dynamic WCF system from the way it is originally presented in Hartshorn (2008). The two instructional methodologies are principally different in the following areas: recycling of writing for both teacher and student, and recycling in relation to manageability; the prompts students responded to; the building of prompt context and establishment of background knowledge; the coding symbols used to mark grammar errors in writing; the allotment of classroom instruction time dedicated to rhetorical functions of writing versus editing grammar mistakes; and the out-of-class learning and assignments that took place.

Each of the aforementioned points will be presented and described in the context of both McKay's and Hartshorn's instructional methodologies, with the purpose of explaining how McKay's instructional methodology is dramatically different from Hartshorn's (2008). Given that the two instructional methodologies are so dissimilar, detailed explanations are warranted.

Recycling. Primarily different between the two methods of instruction was whether or not prompts were recycled (i.e., reused). It was McKay's intention to position this new principle of recycling subordinately within the already-existing dynamic WCF principle of manageability.

Students still responded to four prompts per week as in the Hartshorn (2008) treatment, but the prompts were recycled so that Prompt A was written on both Monday and Wednesday, and Prompt B on both Tuesday and Thursday. The fewer number of new topics introduced was intended to be more manageable for both the student and the teacher, eliminating the need for excessive rewriting and marking of subsequent drafts. An additional aim of recycling was that teachers and students would find the cognitive load more manageable: instead of thinking about four different topics per week, all involved would only have to manage thoughts and ideas revolving around two different, but highly interrelated, topics.

The general flow of writing and rewriting prompts for the two instructional methodologies is outlined in Table 1 and Table 2. In both treatments, students wrote four days per week for ten minutes each day. The difference lies in the fact that prompts were recycled by McKay, but not by Hartshorn (2008). Capital letters in these tables indicate different prompts, and subscript numbers indicate drafts/versions.

Table 1

Weekly Writing in Hartshorn's Traditional Instructional Methodology

Traditional (2008)					
WEEK 1	Mon	Tues	Wed	Thurs	Fri
write	A_1	B_1	C_1	D_1	
edit		A_2	B_2	C_2	
			A_3	B_3	
				A_4	
WEEK 2	Mon	Tues	Wed	Thurs	Fri
write	E_1	F_1	G_1	H_1	
edit	D_2	E_2	F_2	G_2	
	C_3	D_3	E_3	F_3	
	B_4	C_4	D_4	E_4	

In the traditional model, students wrote four prompts per week. They had one week to rewrite a prompt as many times as possible with the objective of achieving error-free writing. Generally students could rewrite prompts up to four times within one week. Therefore, at any given time, students and teacher were involved in writing, marking, rewriting, or editing four different drafts about four different topics. Table 2 shows the modified flow of writing and rewriting prompts, reflecting McKay's goal of making both teacher and student workload more manageable through recycling.

Table 2

Weekly Writing in McKay's Modified Instructional Methodology

	Modified (2013)				
WEEK 1	Mon	Tues	Wed	Thurs	Fri
write	A^1	B^1	A^2	B^2	
rewrite		A^1	B^1	A^2	B^2
WEEK 2	Mon	Tues	Wed	Thurs	Fri
write	C^1	D^1	C^2	D^2	A^3 B^3
rewrite		C^1	D^1	C^2	D^2

Note that in Table 1, subsequent submissions of drafts are referred to as 'edits,' and in Table 2, subsequent drafts are 'rewrites.' This is a critical distinction between the two instructional methodologies and their approaches to draft resubmission which will be discussed later in this work.

In the modified model, students responded to only two prompts per week. On Monday and Wednesday, the same prompt was written; on Tuesday and Thursday, the same prompt was written. The goal was that students would remember the feedback from Monday on Wednesday and the feedback from Tuesday on Thursday, and consequently produce more accurate drafts while experiencing less cognitive overload.

Additionally, students were limited to two rewrites per draft, which were to be turned in on specific due dates. They were given a full week to write a third draft, and expected to utilize dynamic WCF from draft two and therefore produce a third draft entirely free of error. At any given time, students and teacher were involved in writing, marking, rewriting, or editing two drafts about two topics that were related to each other.

In Table 2, it can be seen that third drafts were submitted on Friday, although classes were not held on Friday in either study. This was accomplished by McKay requiring students to place drafts due on Friday in his box, located in the main office of the intensive ESL program building. While students did not attend classes on Friday in either study, they took tests at the school every Friday in both studies. Students were therefore present in the building, and no extra travel or inconvenience was required to submit Friday drafts.

The process of recycling and resultant decreased workload in the modified approach aimed to result in more manageability of the drafting process for both student and teacher, thereby attempting to improve upon one aspect of the four principles of dynamic WCF. Students were not writing or rewriting more than two drafts per day, and the teacher was not marking more than two drafts per day per student. Consider this in comparison to the traditional approach, where students could potentially be rewriting four or more drafts and the teacher marking four drafts times the number of students per day.

Prompts. The prompts students responded to daily for 10 minutes differed substantially between Hartshorn (2008) and McKay. In the traditional method, students responded to short, simple prompts focused on general topics such as “effective leadership,” “care for the elderly,” or “competition.” In the modified method, students responded to more lengthy, multi-sentence official TOEFL prompts from the ETS web site (http://www.ets.org/Media/Tests/TOEFL/pdf/989563wt.pdf) or prompts modeled closely thereafter, such as “Do you agree or disagree with the following statement? Parents are the best teachers. Use specific reasons and examples to support your answer.” Refer to Appendix B for examples of short prompts used in Hartshorn and multi-sentence prompts used by McKay.

Shelley (forthcoming) recommends using two-word prompts such as those in Hartshorn (2008) as the most effective way to elicit student responses. Students are not limited to a narrow topic, particular tense, or set of prompt-driven vocabulary words. Students can write in any style they wish, from personal narrative to expository to compare and contrast. Evans states that the rationale behind short prompts is that if students have only 10 minutes to write, a multi-sentence prompt is likely too much to process. Short prompts target general knowledge: every student should have a thought immediately come to mind, and therefore have something to write about within the first minute of seeing the prompt (personal communication, November 15, 2013). Short prompts were used daily in Hartshorn's study.

In McKay's instructional methodology, multi-sentence TOEFL prompts were intended to limit students in their expression, aimed at eliciting specific structures that may need attention. Students could not write exclusively using grammatical constructions, tenses, or vocabulary words they were comfortable with because the prompts required specific elements and modes of production. Furthermore, students could not employ avoidance strategies, which Truscott (2007) identified as a potential shortcoming of WCF, because prompts targeted the elicitation of specific grammar principles.

Consider, for example, the prompt that follows: "If you could invent something new, what product would you develop? Use specific details to explain why this invention is needed." This prompt is similar to those used in McKay's instructional methodology. In this example, students are forced to use the future unreal conditional and all the modals or auxiliaries associated with it. Under Hartshorn's (2008) traditional instruction, a comparable prompt might have read: "Inventions." McKay believes that a short prompt such as "Inventions" would not

elicit errors that need correction because students would likely avoid the future unreal conditional, favoring tenses or structures that they are more comfortable using.

Evans, on the other hand, believes that multi-sentence prompts may force students to use language that they are not ready or prepared to appropriately use, stating: "Throughout all the time I've been using this method, I never felt like avoidance was an issue." He elaborated further that if students only used language they were comfortable with in writing two-word prompts, drastic improvements in accuracy would not have been possible in Hartshorn (2008) and Evans et al. (2011) since students would never have produced any errors to begin with (personal communication, November 8, 2013).

McKay's primary purpose for switching from short prompts to longer prompts was authenticity: longer prompts are what students would receive on the TOEFL or in a university setting. Thus, the change in prompt type does not reflect one of McKay's modifications targeted at the dynamic WCF principle of manageability, but rather meaningfulness. Short prompts would not be as meaningful as longer, authentic prompts, in McKay's estimation.

Prompt context and background knowledge. The multi-sentence prompts given to students by McKay were coordinated with the students' three other daily classes. The topics were parallel with the content of their other classes (e.g., recycling, environmental issues, biology, and so on). As long as students attended their concurrently enrolled classes, they would have context and background knowledge for the topics about which they wrote in linguistic accuracy and received dynamic WCF.

Moreover, the two prompts written per week in McKay's instructional methodology were intended to be highly related to each other. Prompt A, for example, might read: "Describe either a successful or unsuccessful effort you have seen a government make to conserve natural

resources. Describe what the effort was and why it was a success or failure. Use specific details and examples to support your answer." Prompt B, which would be written the following day, might read: "Do you agree or disagree with the following statement? *One person cannot make a difference in the environment.* Use specific details and examples to support your answer." These two prompts are connected in their general relation to the environment.

The short prompts given to students in Hartshorn (2008) aimed to be more general and without requisite activation of background knowledge. If students saw the prompt "lawyers," for example, it was the mindset in the traditional method that students would have thoughts immediately come to mind from their life experience. Thus, the 10 minutes given would not be used in reading and thinking about a multi-sentence prompt, but rather could be focused entirely on writing accurately and proofreading. Background knowledge presumably already existed without much additional activation, explanation, or thought.

The relatedness of prompts in McKay's method did not occur in Hartshorn's (2008) method. There was no connection between the prompts that students in the Hartshorn study wrote about; thus, they wrote about four different, unrelated topics per week. McKay's attempts to coordinate prompt topics with each other and with students' concurrent classes was a further effort at manageability. Time would not have to be spent activating background knowledge, researching new topics, or preparing presentations on unfamiliar content.

Error markings. The dynamic WCF error markings written by the teacher on student drafts were considerably different between the two instructional methodologies. McKay used markings intended to contain highly specific information for students as they self-corrected.

As an example, teachers in Hartshorn's (2008) study marked an accuracy error in verb form simply as 'VF.' A student who produced *"She should to do her homework"* would receive

a general ‘VF’ marking above the infinitive verb *“to do,”* with no additional or more specific information about the type of verb form error. In McKay’s instructional methodology, a student who produced the same sentence would receive the marking ‘VFb,’ suggesting that *“to do”* should not be the infinitive form of the verb, but the base form. In the same ‘VF’ category, the markings ‘VFi’ verb form infinitive, ‘VFg’ verb form gerund, and ‘VFp’ verb form participle were also used in McKay’s instructional methodology.

Recall back to Ferris (1999), who stated that if error correction was “selective, prioritized, and clear” (p. 4), it would be effective. Indeed, the error markings in both instructional methodologies fulfilled these three objectives. It was McKay’s intention, however, that by modifying the error markings he would fit these criteria to a greater extent than Hartshorn (2008) and make the editing process more approachable and manageable for students. See Appendix A for a full comparison of the two sets of markings.

Classroom instruction. McKay’s daily instruction was similar to that in Hartshorn’s (2008) study, as described in Hartshorn et al. (2010): the “classroom discussions and activities were centered on the most frequent types of errors being produced by the students in their daily writing” (pp. 94-95). There was not a set syllabus with instructional grammar topics, but rather a dynamic syllabus, which was shaped according to diagnostic information and what students produced incorrectly. According to Ferris et al. (2013), “feedback, paired with discussion activities contextualized to the exact problems students are having at that moment, has strong potential to be helpful” (pp. 322-323). Such concentrated discussions, pinpointing the problematic grammar areas as indicated by students’ productions, were a weekly occurrence in both the traditional and modified approaches.

Ferris and Roberts (2001) stated: ". . . survey and interview data from our previous study suggested that students struggled with applying teacher feedback to their writing because they were unfamiliar with the grammatical rules and terminology connected with the . . . error categories on our list" (p. 167). Both Hartshorn (2008) and McKay sought to avoid this pitfall and contextualized error feedback given through grammar-heavy classroom instruction. In both studies, high priority was given to helping students understand grammar terms and practice grammar daily in the classroom so that when WCF was given, students would understand what to do with it and find the error markings meaningful.

Despite those similarities in classroom instruction, McKay's instruction focused substantially more on rhetorical aspects of larger writing tasks than the original instruction did. Unlike instructors in Hartshorn (2008), McKay taught functions of academic writing, essay structure, and pacing for thirty-minute timed writing tasks.

An additional difference lies in how written prompt content was addressed in each instructional methodology. In Hartshorn's (2008) study, students did not modify their four weekly drafts in terms of content. Students received a score out of 10 for each daily written text, weighted 25% for content and 75% for accuracy. Based on that score, students corrected their grammatical errors but were not expected to make modifications beyond editing grammar. They were not taught rhetorical writing functions during classroom instruction time and were therefore not expected to incorporate such knowledge into their daily prompt writing.

In contrast, McKay devoted a much greater focus to expanding paragraph-level writing to essay-level writing and provided instruction on this during class time. Students were expected to modify or expand their drafts in terms of content, and not only edit them for grammatical accuracy. In the current study, students' third and final draft was an expanded essay of

approximately one full page with an introduction, two brief body paragraphs, and conclusion that was not only edited for grammatical accuracy but also expanded in content. Students were encouraged to include certain functions of academic writing in these expanded drafts, such as a counterargument and topic sentences.

As it was originally conceptualized by its developers, dynamic WCF should encourage teachers: "…Push your students to focus on EDITING, not rewriting. This isn't a composition class. …Discourage them from making any changes that aren't intended to correct the grammar errors marked" (Shelley, forthcoming, p. xx). McKay dramatically modified dynamic WCF in his adherence to this recommendation by focusing not only on editing, but on non-dichotomous classroom instruction addressing both composition and grammatical editing.

Out-of-class learning and assignments. In Hartshorn (2008), students kept an error tally sheet, error list, and edit log. The error tally sheet was simply a numerical record of errors marked by the teacher on each draft. The error list consisted of a heading for each error (e.g., determiners), under which students copied all of their sentences containing that error. The edit log was a record of how many rewrites a student produced before the draft was deemed free of error. See Appendix C, D, and E for examples of Hartshorn's three out-of-class assignments.

There is strong evidence for a rationale behind students' creation of an error list and error tally sheet. The students in Hartshorn (2008) saw great improvements in their accuracy as they carefully maintained lists and tallies of their errors. Polio (2012) declared that correcting errors

> seems essentially useless if learners do not have anything to do with the feedback. What [teachers] need to consider is whether or not there is a better way to draw learners' attention to the corrections other than simply rewriting. Some teachers have students keep logs of their error types, for example. (p. 385)

Guénette (2007), advocating WCF, stated "the fact that positive results were seen in the short term shows that pedagogical intervention that pushes learners to pay attention to the language is useful" (p. 44).

Similar to the instructors in Hartshorn's (2008) study, McKay required students to keep an error tally sheet and error list; however, an edit log was not required. The edit log was replaced by what the instructor named an Academic Input Inventory. For this assignment, students were required to collect ten academic words or phrases from credible sources. These were items that they could incorporate into their prompt the next time they rewrote it to improve academic content and word choice. Its purpose was intended to develop the content of writing more than the grammatical accuracy of writing, which generally separates it from anything that would have been done in the instructional methodology of Hartshorn.

Another point of interest is that McKay modified the error list assignment from the way it was implemented in Hartshorn (2008). McKay did not require students to write down every single sentence containing an error. Rather, he required students to choose ten sentences reflective of their major errors, and write them down both as they were incorrectly produced and then as they would be written correctly. It was McKay's aim that writing ten sentences on the error list would be more manageable than writing all sentences as required in Hartshorn's method. See Appendix F, G, and H for examples of the three out-of-class assignments required by McKay.

Summary of similarities and differences. Table 3 provides a description of similarities between the traditional (2008) and modified instructional methodologies. In Table 4, a comprehensive summary of the differences between the traditional (2008) and modified instructional methodologies is given.

Table 3

Summary of Similarities between the Two Instructional Methodologies

Aspect of dynamic WCF	Description
Timely	Feedback given the next day
Constant	Students wrote for 10 minutes every day, four days a week
Meaningful	Students were taught coded error markings and the grammar principles behind them

Table 4

Summary of Differences between the Two Instructional Methodologies

	Instructional Methodology	
Aspect of dynamic WCF	Traditional, Hartshorn (2008)	Modified, McKay (2013)
Recycling	Prompts never reused in class	Same prompt written twice in class
Manageability	Teacher involved in marking up to four drafts per student per day	Teacher involved in marking up to two drafts per student per day
10-minute writing prompts	Short prompts allowing freedom of expression, not aiming to elicit certain grammar constructions	Multi-sentence, restrictive TOEFL prompts aimed at eliciting certain grammar constructions
Establishing prompt context and background knowledge	General topics of common knowledge; activation of background knowledge not required	Specific topics coordinated with each other and concurrent classes
Marking symbols	More general, identifying broad error categories	More specific, identifying narrow error categories
Classroom instruction	Focused only on grammar	Focused on both grammar and rhetorical aspects of writing
Drafting process	Subsequent drafts focused on editing grammar mistakes only	Subsequent drafts focused both editing grammar mistakes and modifying content
Out-of-class learning	Error tally sheet; error list; edit log	Error tally sheet; error list (modified); Academic Input Inventory

While a wide variety of factors were dramatically different as implemented by McKay (recycling as a subset of manageability; the prompts students responded to; the building of prompt context and establishment of background knowledge; the coding symbols used to mark grammar errors in writing; the allotment of classroom instruction time dedicated to rhetorical functions of writing versus editing grammar mistakes; and out-of-class learning and assignments), what was common between the two instructional methodologies was their adherence to the dynamic WCF principles of timely, constant, and meaningful. In spite of their dramatic differences, both approaches to dynamic WCF sought to preserve the four key principles and center all instructional practices around them.

Eliciting and Analyzing Data

Although McKay's model of dynamic WCF is dramatically divergent from the model as originally conceptualized, similar elicitation procedures and data analysis procedures were implemented to keep the research designs as consistent as possible and answer the question of what effect the dramatic modifications would have on students' L2 writing accuracy.

Elicitation procedures: pretest and posttest. As part of classroom instruction under both Hartshorn (2008) and McKay, students were required to write several 30-minute essays throughout the course. Similar to Hartshorn, the complexity, accuracy, and fluency of a pretest essay at the beginning of the semester and a posttest essay at the end of the semester were used as data in the current study.

The pretest and posttest prompts in the current study were identical to those used in Hartshorn's (2008) study, with the goal being to replicate the original research design as closely as possible. At the beginning of the semester, students responded to the following prompt for a pretest: "Do you agree or disagree with the following statement? Only people who earn a lot of

money are successful. Use specific reasons and examples to support your answer." The posttest prompt at the end of the semester read: "In your opinion, what is the most important characteristic (for example, honesty, intelligence, a sense of humor) that a person can have to be successful in life? Use specific reasons and examples from your experience to explain your answer."

Anonymity of data. The pretest and posttest essays written by students at the beginning and end of the semester under study were randomly assigned a seven-digit identification number by the testing administrators in the instructional context. Raters analyzing the data never saw a student name associated with a piece of writing, and the seven-digit identification numbers had no relationship whatsoever to the student who produced a piece of writing.

Analyzing data. Three raters analyzed the pretest and posttest essays (R1, R2, and R3). To begin, R1 and R2 individually divided the pretests and posttests into T-units, and afterwards worked jointly to confirm the correctness of all T-unit divisions. Where there was a discrepancy, R1 and R2 reexamined the T-units and reached an agreement. T-units were primarily chosen for use in this study in an attempt to replicate as closely as possible Hartshorn (2008), since he used the same unit of analysis. Hunt (1965) originally developed the concept of a T-unit, and defined it as consisting of "one main clause plus the subordinate clauses attached to or embedded within it" (p. 49). After division into T-units, the pretests and posttests were then distributed to R1, R2, and R3 for analysis in three major areas: complexity, accuracy, and fluency.

Complexity. For this study, complexity was defined as the average number of words per unit of expression (e.g., sentences; clauses; T-units) in an essay. As in Hartshorn (2008), T-units were the chosen division to create these units of measurement for complexity analysis in the current study. Thus, the measurement of mean length of T-unit (MLTU) was used to determine

complexity on pretest and posttest essays in both studies. To calculate the MLTU, R1 divided the word count of each essay by the total number of T-units.

Accuracy. Wolfe-Quintero, Inagaki, and Kim (1998) recommended EFT/T to determine writing accuracy. EFT/T is the total number of error-free T-units (EFT) divided by the total number of T-units (T) present in the writing" (p. 90). Since Hartshorn (2008) used the EFT/T ratio to determine accuracy, the current study did the same.

The T-units in all pretest and posttest essays were marked individually by R1, R2, and R3 as being either "errored" or "error-free." R1 received reports from all three raters on their number of T-units judged as error-free for each of the 48 essays. R1 calculated the EFT/T for the 48 essays for each of the three raters and input the ratios into a spreadsheet. Although there was not 100% agreement on the EFT/T ratios for each essay, inter-rater reliability was high with an intraclass correlation coefficient of .968 among the three raters.

Fluency. Hartshorn (2008) references Wolfe-Quintero et al. (1998) for a definition of fluency: "a measure of the sheer number of words or structural units a writer is able to include in their writing within a particular period of time" (p. 45). In the current study, the word count of each essay was determined by a word processor and recorded in a spreadsheet by R1.

Although it was examined in Hartshorn (2008), rhetorical competence was not within the scope of the current study. This work focuses on accuracy and a manipulation of dynamic WCF principles aimed at improving accuracy. R1, R2, and R3 informally judged rhetorical competence in the current study by providing approval for each essay that it addressed the topic of the given prompt. Relevancy to the prompt topic occurred 100% of the time among the 48 pretest and posttest essays, and thus will not be addressed in the following section.

Results

To contextualize the results, recall the two research questions of this study: (1) Based on 30-min pretest and posttest essays, will mean accuracy scores in a group experiencing a dynamic WCF treatment dramatically modified from Hartshorn (2008) increase, decrease, or remain the same as scores obtained by students in the Hartshorn study? (2) Based on 30-min pretest and posttest essays, will there be significant compromises in fluency or complexity?

This section presents results with the objective of answering these two research questions. Reliability estimates and information about pretest scores are given, and complexity, accuracy, and fluency scores between students experiencing McKay's instruction and students who experienced Hartshorn's instruction are compared.

Reliability Estimates

The Pearson correlation coefficients between R1 and R2, R1 and R3, and R2 and R3 were 0.945, 0.934, and 0.913 respectively. Hartshorn (2008), in describing the Pearson correlation coefficient, states that "higher values . . . indicate greater strength in the linear relationship between two raters" (p. 70). This statistic can account for the degree of reliability between two raters. An intraclass correlation coefficient, on the other hand, was used by Hartshorn because it "provided an average measure of consistency among all three raters" (p. 72). In other words, this statistic can account for more than two raters and thus is important to note in addition to the three Pearson correlation coefficients. The intraclass correlation coefficient among all three raters for agreement on error-free T-unit ratios in the current study was .968. This statistic suggests that there was a high degree of reliability among the three raters, and that their judgments regarding T-unit accuracy were reliable.

Two Treatment Groups: Two Starting Points

The treatment group in Hartshorn's (2008) study (n = 28) and the students experiencing McKay's instruction in this study (n = 24) began their semesters at different levels of complexity, accuracy, and fluency. In other words, the pretest scores of the two treatment groups show variance worth noting. Generally speaking, McKay's students scored higher at the onset in their complexity, accuracy, and fluency than Hartshorn's students, suggesting that the McKay's group was slightly more advanced to begin with.

On the pretest essay at the beginning of the treatment, McKay's students averaged 16.03 words for MLTU, while Hartshorn's (2008) students averaged 13.71. The slight difference between the beginning complexity levels of the two treatment groups is evident. Similarly, regarding accuracy, McKay's students produced an average ratio of .199 EFT/T, while Hartshorn's students averaged .141 EFT/T. Interestingly, while the two groups did differ in their fluency starting points, fluency is where the difference was the most minor. The current group averaged 359.17 words per pretest essay, and the Hartshorn group averaged 357.36 words.

Complexity, Accuracy, and Fluency Results

The results of the current study and the effects of McKay's dramatically modified instructional methodology on students' complexity, accuracy, and fluency in L2 writing will be presented. Complexity, accuracy, and fluency are common across the literature as measures for analyzing L2 writing (Evans et al., 2010, 2011; Hartshorn, 2008; Hartshorn et al., 2010; Hartshorn & Evans, 2013) and constitute a prominent set of descriptors about L2 writing (Skehan, 1998). Hartshorn (2008) states: "In addition to linguistic accuracy . . ., writing fluency and writing complexity are also commonly used by researchers to measure writing development." While fluency and complexity are secondary to accuracy in answering the

research questions, as Hartshorn states, “including them [will] help contextualize findings and expose possible unintended consequences of the treatment on L2 writing production” (p. 36).

Statistics will be presented in several formats and will convey information aimed to answer the research questions. First, the pretest and posttest mean and standard deviation values will be presented for both Hartshorn (2008) and McKay independently. Next, the results of an ANOVA test comparing the two groups will be given—i.e., statistical significance values (p) and effect sizes (η_p^2) resultant of comparison between McKay’s and Hartshorn’s instructional methodologies. Third, independent statistics void of comparison will be presented. Paired sample t-tests were run to produce these results, which convey information about the instructional methodologies independent of each other and their effects on the three areas of L2 writing analyzed. In describing the paired sample t-test results, p will be identified as well as Cohen’s d to represent effect size. Finally, figures will be presented which provide graphical representation comparing the two groups.

In Hartshorn (2008), only mixed model, repeated measures Analysis of Variance (ANOVA) tests were carried out to produce the statistical results. Since an analysis of the interaction of multiple variables was necessary, an ANOVA test was appropriate. In the current study, however, both ANOVA and paired sample t-tests were run by researchers. ANOVA tests were used to produce a statistical significance value of McKay’s modified instructional methodology in comparison to Hartshorn’s, and paired sample t-tests were used to show the statistical significance of McKay’s and Hartshorn’s instructional methodologies independent of each other. Given the single treatment group of the current study, paired sample t-tests were deemed more reasonable to make McKay’s and Hartshorn’s independent data completely comparable.

Of particular interest in the ANOVA (comparison) statistics are the *p*-values (*p*), representing statistical significance, and the partial eta squared values (η_p^2), representing effect size or practical significance. In the current study, a *p* below .01 is considered statistically significant; therefore, a *p* above .01 is non-statistically significant. Additionally, the η_p^2 values, or effect sizes, are important to note in conjunction with *p* because "the results of some research may be statistically significant while practical significance is negligible" (Hartshorn, 2008, p. 93). The η_p^2 provides a reflection of practical significance and measures how large or small of an impact one variable had on another. Hartshorn references Cohen (1988) and his guidelines for interpreting η_p^2 statistics: .01 represents a small effect, .06 represents a moderate effect, and .14 or above represents a large effect (p. 96, 112).

Hartshorn (2008) calculated η_p^2 as opposed to eta squared (η^2), and provides an extensive description of the differences between these two measures of practical significance or effect size. Essentially, η^2 is

> flawed in that the strength of association depends on how many independent variables are included in the design and how significant those variables are. Thus, the reliability of the η^2 statistic as an estimate of effect size seems somewhat context dependent. (p. 95)

While η_p^2 has its own limitations, it is generally considered more useful by researchers. In summarizing Bakeman and Robinson's (2005) explanation, Hartshorn writes that "unlike η^2, the η_p^2 is rather successful at isolating the effect of a specific variable" (p. 96). Thus, η_p^2 is used in the current study to measure the effect size of McKay's dynamic WCF instructional methodology on complexity, accuracy, and fluency when compared to Hartshorn.

Of particular interest in the paired sample t-test (independent) statistics are *p* and Cohen's *d*, representing effect size. As in the ANOVA tests, *p* was set at .01 by the current researchers to

indicate statistical significance. Cohen's *d* represents effect size, or practical significance. According to Cohen (1988), *d* = .10 is considered a small effect size, *d* = .30 is considered a medium effect size, and *d* = .50 is suggested to indicate a large effect size or high practical significance. Other numbers, including sums of squares (*SS*), degrees of freedom (*df*), mean squares (*MS*), and F-statistics (*F*), will not be discussed as they are not relevant to answering the research questions of this study.

Complexity. The effect of McKay's instructional methodology on L2 writing students' complexity was negligible. The average MLTU in the current study was 16.03 words at the beginning of the semester and 15.32 words at the end of the semester, showing that students experienced only a small decrease of 0.71 words per T-unit in their complexity. In Hartshorn (2008), the students started at a lower level of complexity with an average MLTU of 13.71 words on the pretest and decreased 0.12 words per T-unit on the posttest, thus declining to an average MLTU of 13.59 words. While Hartshorn's students experienced less of a complexity decrease than McKay's group, Hartshorn's group showed a lower level of complexity overall on both the pretest and posttest. Table 5 summarizes the pretest and posttest statistics for the two treatment groups.

Table 5

Descriptive Statistics for Writing Complexity

Group		Pretest	Posttest	Decrease
McKay (2013) (n = 24)	Mean SD	16.03 2.56	15.32 2.91	-0.71
Hartshorn (2008) (n = 28)	Mean SD	13.71 2.49	13.59 2.48	-0.12
Average (N = 52)	Mean SD	14.78 2.75	14.39 2.8	-0.39

These results are statistically contextualized by ANOVA and paired sample t-tests. When compared to Hartshorn's (2008) treatment's effect on complexity, McKay's treatment was not statistically significant and had a small effect (p = .448, η_p^2 = .012). Additionally, independent statistics show that complexity was not significant in either McKay's modified treatment (p = .222, d = .256) or Hartshorn's traditional treatment (p = .821, d = .043).

Figure 1 provides a graphical representation of complexity decreases in both Hartshorn's (2008) treatment group and McKay's current group. It is evident that while the two groups had different starting points, both experienced slight declines in overall complexity from pretest to posttest.

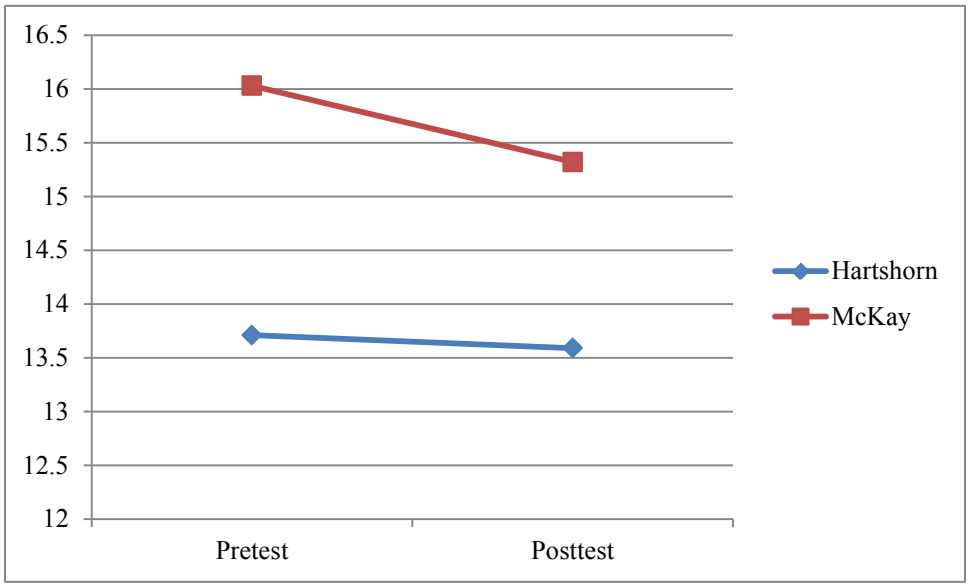


Figure 1. Pretest and posttest means for writing complexity

It can be inferred from the data that there was not a great difference between the effects of McKay's and Hartshorn's (2008) instructional methodologies on posttest complexity—in both studies, complexity was not statistically significant, the effect sizes of the two instructional methodologies on complexity scores were small, and students experienced negligible decreases in MLTU. It is interesting to note, however, that McKay's numbers come closer to both

statistical and practical significance (p = .222, d = .256) than Hartshorn's numbers (p = .821, d = .043), suggesting that perhaps McKay's treatment impacted complexity slightly more negatively.

Accuracy. The L2 writing students receiving McKay's instruction in the current study experienced an average EFT/T ratio increase of .105 from pretest to posttest essay. McKay's students started out with EFT/T ratios of .199 on average per essay, and ended up at .304 EFT/T on the posttest at the end of the semester. This is comparable to Hartshorn (2008), in which students experienced an average increase of .102 in EFT/T. Hartshorn's students started out at .141 EFT/T and increased to .243 by the end of the semester. This information is summarized in Table 6.

Table 6

Descriptive Statistics for Accuracy Scores

Group		Pretest	Posttest	Increase
McKay (2013) (n = 24)	Mean	.199	.304	.105
	SD	.138	.199	
Hartshorn (2008) (n = 28)	Mean	.141	.243	.102
	SD	.149	.194	
Average (N = 52)	Mean	.168	.271	.1035
	SD	.146	.197	

When compared to Hartshorn's (2008) instructional methodology, McKay's instruction did not differ statistically or practically (p = .932, η_p^2 < .001) in its effect on L2 writing accuracy. Independently, however, both instructional methodologies were highly significant (p < .001) in increasing L2 writing accuracy from pretest to posttest. Furthermore, both McKay's and Hartshorn's methods of instruction showed large effect sizes of d = .881 and d = .861

respectively, suggesting that the two methods were essentially equivalent in their effects on accuracy.

In Figure 2, noticeable increases in T-unit accuracy can be seen with both McKay's and Hartshorn's (2008) treatment groups. The sharp incline in EFT/T from pretest to posttest is virtually parallel between the two groups, in spite of differing starting points and instructional methodologies experienced.

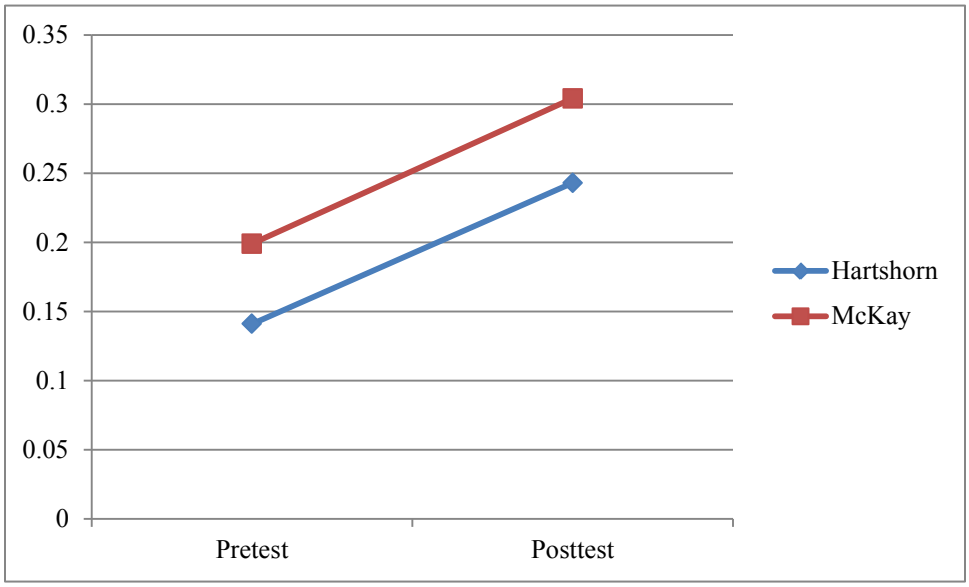


Figure 2. Pretest and posttest means for accuracy scores

Because of the statistical significance of both McKay's and Hartshorn's (2008) instructional methodologies, and based on the nearly parallel linear relationship between the two groups in Figure 2, it can be reasonably concluded that McKay's approach to dynamic WCF did not affect students' posttest accuracy scores in any way, negatively or positively, when compared to Hartshorn's approach to dynamic WCF. Students receiving McKay's instruction in the current study experienced gains in accuracy comparable to Hartshorn's students.

Fluency. In Hartshorn (2008), the control group of students receiving traditional process writing instruction "increased their fluency significantly more than the students in the treatment

group" (p. 103), with average gains of 50 words per essay compared to only 15 words per essay for the treatment group. Hartshorn's control group, which showed statistically more significant improvements in fluency ($p < .001$) than the treatment group ($p = .03$), wrote an average of 35 more words posttest per essay than the treatment group.

These control group fluency statistics from Hartshorn (2008), however, are substantially less than the approximate 122.96-word gain experienced by students in the current study under McKay's instruction, who leapt from producing an average of 359.17 words in their pretest essays to 482.13 words in their posttest essays. Students receiving McKay's instruction produced approximately 73 more words on average on the posttest than students in Hartshorn's control group who specifically received traditional process writing instruction.

Table 7 provides information comparing writing fluency for McKay's current treatment group and Hartshorn's (2008) treatment group. Of particular interest is that McKay's students wrote approximately 108 more words per posttest essay than students in Hartshorn's treatment.

Table 7

Descriptive Statistics for Writing Fluency

Group		Pretest	Posttest	Gain
McKay (2013) (n = 24)	Mean	359.17	482.13	122.96
	SD	66.39	106.33	
Hartshorn (2008) (n = 28)	Mean	357.36	372.75	15.39
	SD	89.08	117.19	
Average (N = 52)	Mean	358.19	423.23	65.04
	SD	78.68	124.1	

To statistically contextualize these results, an ANOVA test comparing the two treatments showed that McKay's instructional methodology had a large effect ($\eta_p^2 = .295$) on fluency compared to Hartshorn's (2008) instructional methodology, and McKay's instruction was

statistically significant ($p < .001$) in improving fluency. Paired sample t-tests showed that independently, McKay's instruction improved fluency from pretest to posttest with high significance and a large effect ($p < .001$, $d = 1.532$), while Hartshorn's posttest fluency results were not significant ($p = .363$, $d = .175$).

It is apparent in Figure 3 that McKay's students experienced a sharp increase in fluency compared to Hartshorn's (2008) students, increasing their writing by 122.96 words compared to Hartshorn's students' 15.39 words.

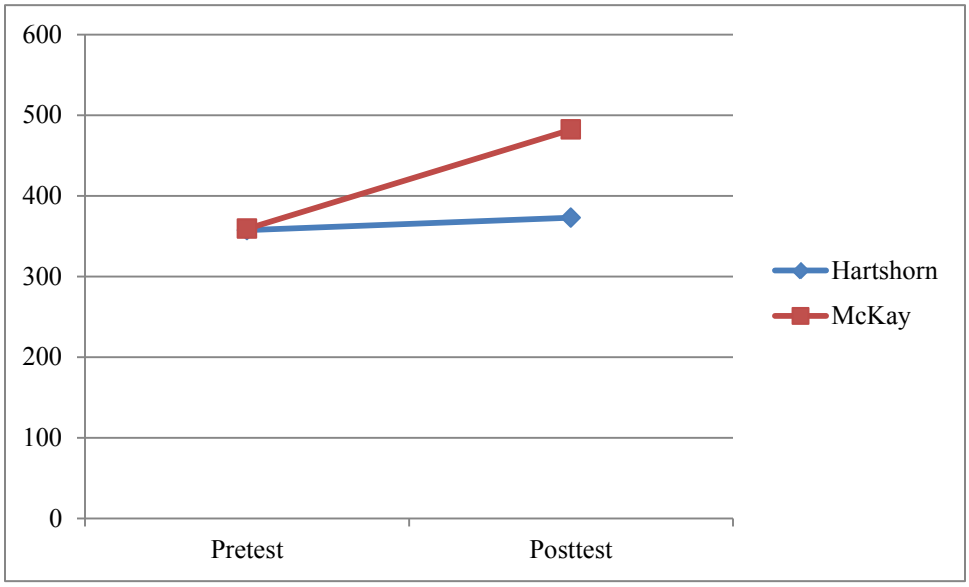


Figure 3. Pretest and posttest means for writing fluency

In summarizing the fluency results, it is apparent that some internal or external variable affected McKay's students' fluency drastically in comparison with Hartshorn (2008), and that this same variable likely did not exist in Hartshorn's study. Further examination into this phenomenon will take place later in this work. While McKay's instructional methodology had a large effect size on his students' posttest fluency scores with an accompanying high level of statistical significance, Hartshorn's data shows his instructional methodology had no effect size and was not statistically significant.

Rhetorical Competence

As was previously explained, R1, R2, and R3 did not extensively judge the rhetorical competence of the pretests and posttests in this study or use a rubric to assign a rhetorical competence score to each essay. This is because the current research, as addressed by the research questions, was primarily concerned with McKay's modified instructional methodology's effects on accuracy and whether the modifications made would increase or decrease accuracy scores.

While Hartshorn's (2008) raters used a writing rubric adapted from the TOEFL iBT and assigned each piece of writing a rhetorical competence score between 0 and 5, the current study's raters simply judged whether each piece of writing was on-topic or off-topic in relation to the prompt. All three raters found 100% appropriateness to the prompt across all pretest and posttest essays; in other words, none of the 48 pretest or posttest essays addressed a topic different than the one presented by the prompt.

Conclusions

Following is a discussion which elaborates on the results presented in the previous section. The data results for complexity, accuracy, and fluency are described and made more meaningful by comparison to Hartshorn's (2008) data. Current data is pedagogically contextualized in terms of application for practitioners and instructors. Pedagogical implications are addressed, as well as limitations of the current study and possibilities for future research.

Pedagogical Implications

While some researchers believe, regarding WCF, that there is "no value for the practice" or that "clearly, grammar correction is not effective" (Truscott, 1996, p. 341), the current study seems to echo Hartshorn (2008) in suggesting otherwise and adds to a growing body of research on the efficacy of dynamic WCF as a practice for improving accuracy in L2 student writing (Evans et al., 2010, 2011; Hartshorn et al., 2010; Hartshorn & Evans, 2013).

Even when drastically modified or implemented in a different way, dynamic WCF as a practice for error correction has positive effects on L2 writing accuracy. The results of this study confirm previous research in suggesting that dynamic WCF can be effective in some contexts and with advanced ESL students. As long as the core of the instructional methodology adheres to the four principles of timely, constant, meaningful, and manageable, and an instructor follows a dynamic syllabus based on student-produced errors, advanced-level students will likely see gains in linguistic accuracy.

Complexity. There were negligible effects on the complexity of the writing of L2 students who experienced McKay's dramatically modified treatment. Students in both studies decreased their MLTU by <1 word, specifically 0.71 under McKay's instruction and 0.12 under Hartshorn's. Instructors practicing dynamic WCF in any form may see slight decreases in their

students' complexity, but this is not certain. Since classroom instruction should focus heavily on accurate production of grammatical forms, it is possible that the lack of attention given to sentence structure results in a decrease of complexity.

The slight decrease in complexity that students experienced in the current study is not far distant from the slight decrease experienced by students in the Hartshorn (2008) study, and is not significant (p = .222) with a low effect (d = .256). It may be inferred that there is indeed a delicate relationship among the three elements of complexity, accuracy, and fluency, as claimed by Skehan (1998). Changes in one component of the complexity, accuracy, and fluency model will not leave other components unaffected. Results of this study, regardless of inconsequentiality, confirm Skehan's notion that gains in one area may lead to compromises in another. In the instance of this study, a shorter MLTU resulted in higher accuracy; in other words, sacrifices in complexity contributed to positive changes in accuracy.

Accuracy. McKay's treatment had no effect on accuracy scores, adverse or positive, in comparison with the effect of Hartshorn's (2008) treatment (p = .932, η_p^2 < .001). McKay's students increased their writing accuracy EFT/T ratios by .105 and Hartshorn's students increased by .102, a difference of merely .003. Furthermore, since both McKay's and Hartshorn's instructional methodologies were statistically significant (p < .001), the treatments may have been essentially equivalent in their influence on L2 writing accuracy.

Based on the results of the current study, it can be concluded that the dramatic modifications McKay imposed (recycling; prompts; prompt context and background knowledge; error markings; allotment of classroom time; out-of-class learning and assignments) did not negatively affect accuracy. By the same token, the modifications also did not positively affect accuracy to any greater extent than Hartshorn's (2008) original methodology did. This null

effect, however, is not entirely undesirable. Recall that this study hoped to contribute to an understanding of the question posed by Evans et al. (2010) regarding the *how* of effective error correction in relation to accuracy increase: "the more essential question is *how* we help our students write more accurately" (p. 2). Indeed, both methods appear to help students improve their L2 writing accuracy and can be considered effective approaches to error correction; thus, dynamic WCF in either form contributes to our understanding of the *how* question.

Fluency. McKay's students in the current study experienced an unanticipated outcome and possible side effect of McKay's modified instructional methodology: markedly increased fluency. Chandler (2003) noted that in a control/treatment-group research design, "both groups increased significantly in fluency…and that there was no significant difference between the two groups in this improvement" (p. 279). This is quite opposite of the current study, in which McKay's group displayed marked gains in fluency when compared to Hartshorn (2008).

The data of the current study indicates that there were statistically significant increases in McKay's students' fluency ($p < .001$). Students increased the word count of their essays at an average of 122.96 words per essay. Additionally, $d = 1.532$ suggests a large effect size that may be attributed to the instructional methodology. Students experiencing Hartshorn's (2008) treatment increased their fluency by only 15.39 words per essay, substantially less than students receiving McKay's treatment, with no significance or notable effect size ($p = .363$, $d = .175$).

There are two possible reasons why McKay's students in the current treatment group could have experienced such a significant fluency increase. The first reason could be attributed to the fact that McKay's instruction focused on rhetorical aspects of writing and functions of timed essay writing in addition to grammar, while Hartshorn's (2008) instruction did not. McKay's grammar instruction was paired with writing instruction, and thus the class was not

solely an applied grammar class as it was for Hartshorn's treatment group. The combination of dynamic WCF addressing grammar in the students' L2 writing and explicit training and feedback on rhetorical functions and organization of 30-minute essays may have contributed to the fluency increase of McKay's students in the current study.

A second possible reason explaining the fluency explosion in the current study is that McKay's students were concurrently enrolled in an intensive process writing class. In Hartshorn's (2008) study, students were enrolled in only one class or the other: the control group was enrolled in a writing class focusing on traditional process writing instruction; the treatment group was enrolled in a linguistic accuracy class employing dynamic WCF. In other words, the control group did not receive dynamic WCF and the treatment group did not receive traditional process writing instruction. In the current study, however, McKay's students were enrolled in both a writing class focusing on traditional process writing, and a separate linguistic accuracy class focusing on grammar in writing and dynamic WCF. Thus, it is possible that the fluency increase in the current study is independent of McKay's instruction, but could be attributed to instruction students received or practice they had producing longer pieces of text in their concurrent traditional process writing class.

Practical Recommendations

As noted, this study did not provide conclusive evidence to answer the question of "how to use WCF to maximize ESL student opportunities to learn to improve the linguistic accuracy of their writing" (Hartshorn et al., 2010, p. 85). The two instructional methodologies, although drastically different, appear to have had a similar impact on L2 students' accuracy. McKay's students were neither helped nor hurt by the modified instructional methodology. Thus, the current study can only lend to an understanding of dynamic WCF as an effective overall

approach to error correction in L2 writing and validate previous research findings, since neither Hartshorn's (2008) nor McKay's instructional methodology stands out as superior in improving accuracy. Indeed, as Celce-Murcia (2013) suggests, there is no singular best method.

In light of this information, practitioners in intensive ESL programs with the objective of increasing their L2 students' writing accuracy may want to consider employing dynamic WCF. While it is true that there was no observable increase in McKay's students' accuracy compared to Hartshorn's (2008) students, a similar improvement took place nonetheless. This suggests that dynamic WCF is effective so long as it adheres to the four core principles of timely, manageable, constant, and meaningful. The primary constraint is that dynamic WCF has been proven effective only at advanced levels (Lee, 2009) and thus may not be appropriate in lower-level contexts.

The fundamental question remains: Why were fluency scores so drastically different between McKay's and Hartshorn's treatment groups? Practitioners who decide to employ dynamic WCF in their L2 writing accuracy classrooms should be aware that outcomes of their students' fluency are not predictable, and that variations in McKay's instructional methodology were likely not the conclusive reason for the substantial fluency increase.

Limitations

Many limitations existed in the current study, and addressing them is requisite to create a clear and complete picture of this study. The limitations include a multiplicity of variables, differences from Hartshorn (2008) in research design, the enrollment of students in a concurrent traditional process writing class, and the use of T-units in analyzing complexity and accuracy. These limitations will be fully addressed in the subsequent sections.

Multiplicity of variables. In the current study, McKay's instructional methodology underwent numerous, dramatic modifications from the original instructional methodology as conceptualized by Hartshorn (2008). Too many changes were made all at the same time, so it is impossible to isolate one variable and determine its effect on L2 writing accuracy. McKay essentially took the dynamic WCF method as envisioned and employed by Hartshorn and changed it in every aspect except adherence to the principles of timely, constant, and meaningful. It is therefore difficult to draw any conclusions about the results because so many variables were manipulated concurrently.

Research design. There were several inconsistencies in research design between the current study and Hartshorn (2008). First, Hartshorn's research took place with various instructors and groups of students. The 28 students in Hartshorn's treatment group were divided into several smaller classes, each having a different instructor. In contrast, McKay's treatment occurred with one instructor and one group of students. McKay was the sole instructor and taught all 24 students together in the same classroom.

One might assume that McKay's large class size (n = 24) would have detrimental effects on students' overall performance or motivation. Indeed, numerous students indicated on a post-semester questionnaire for McKay's linguistic accuracy class that they prefer smaller class sizes and did not enjoy having 24 students in the same classroom with only one instructor. However, the fact that improvements in accuracy were virtually unaffected between Hartshorn's (2008) students (organized into smaller classes) and McKay's students (placed into one big class) suggests that a large class size is possibly not a variable worth factoring into the efficacy of dynamic WCF.

Second, in Hartshorn (2008), dynamic WCF was novel in the intensive English language program where his study took place. It had never before been employed, and students were unfamiliar with the drafting process and set of error markings. In contrast, students in McKay's classroom had gone through dynamic WCF once or twice before and were highly familiar with the drafting process and error markings generally. This introduces the possibility of some kind of ceiling effect: could there be a saturation point or upper limit at which dynamic WCF is no longer as successful?

Finally, the students in the current study started at a slightly higher level of complexity, accuracy, and fluency overall than students in Hartshorn's (2008) study, according to pretest scores. In other words, although all students in both studies were placed at ACTFL (2012) advanced levels by an institutional placement test, there is a possibility that McKay's students were at a somewhat higher level to begin with.

Concurrent writing class. Not only were there many internal factors that were different from Hartshorn (2008), but a noteworthy external factor as well. Students receiving McKay's instruction were concurrently enrolled in a traditional process writing class, which differs from students receiving Hartshorn's instruction, who were not enrolled in a concurrent writing class. The consequence of this is an unclear reason for why McKay's students' fluency changed so drastically, as it is possible that this occurred because of the concurrent writing class.

Unit of analysis: T-unit. Because this study sought to replicate Hartshorn (2008) as closely as possible, T-units were chosen for complexity and accuracy analyses. However, other units of measurement might be more effective. Evans et al. (2011), for example, rejected T-units, instead using clauses as the unit of analysis in a study on the accuracy of U.S. university-

matriculated L2 students receiving dynamic WCF. The researchers state that clauses were used because they are "believed to be more discriminating than the error-free T-unit ratio" (p. 237).

Evans, Hartshorn, Cox, and De Jel (forthcoming) identify several potential shortcomings of T-units: they introduce the need for reliability statistics and researchers don't always report these in the literature; they do not "account for error severity . . . very minor errors carry the same weight as much more consequential errors" (p. xx); and because they are very long (one T-unit can have multiple clauses), an essay could potentially be marked as zero percent accurate.

In an investigation of alternatives to T-units, Evans et al. (forthcoming) examined weighted clause ratios (WCR). The WCR was developed by Wigglesworth and Foster (2008), who believe that "while an effective measure of linguistic accuracy is vital for research, all current approaches have been problematic . . . without a more precise measurement, potential differences in accuracy may be missed" (Evans et al., forthcoming, p. xx). Wigglesworth and Foster's solution to this problem is WCR, "a means of analysis where each clause is assigned a weight based on the ease of retrieving meaning" (p. xx). Essentially, a text is divided into clauses and assigned one of four weights, based on the assumption that "inaccuracies in units of language will affect comprehensibility to varying degrees" (p. xx).

The results of Evans et al. (forthcoming) show that WCR has high construct validity and correctly represented the accuracy as well as the communicative adequacy of writing samples. The researchers concluded that

> such a measure of accuracy could be much more precise than other measures currently available to researchers . . . because of its rationale and the way it was designed to overcome the limitation of other measures, we find this approach to measuring linguistic accuracy compelling (p. xx).

A limitation of the current study could therefore be that it employs T-units as the unit of analysis. Perhaps, as forthcoming literature suggests, there are better approaches to measuring linguistic accuracy.

Future Research

Using a similar research design and in a comparable intensive English language context with advanced students, future researchers could expand on dynamic WCF research in several ways. Researchers could manipulate only one aspect of dynamic WCF at a time, use a unit of analysis different from T-units, or further pursue the *how* question of optimal error correction by focusing more specifically on manageability and practicality from a practitioner's perspective. These possibilities are explored in greater detail in the following sections.

Isolate one variable. Because so many elements of the dynamic WCF model were changed all at once by McKay in the current study, it was difficult to make any causal inferences. Future studies could change only one element of the instructional methodology and measures its effect on L2 writing accuracy. This would lead to a clearer view of the individual variables involved in dynamic WCF and how manipulation of one of them might change the outcome or affect student improvement. For example, a future study could require students to write only three prompts per week instead of four, and change absolutely nothing else about the instructional methodology. A more conclusive discussion could then be had about how manipulation of one such variable affected L2 writing accuracy.

Use a different unit of analysis. Recent research has investigated the usefulness of various units of measure to analyze L2 writing. Evans et al. (2011) used error-free clause ratios (EFCR), and Wigglesworth and Foster (2008) suggest WCR (as cited in Evans et al., forthcoming). Researchers in a future study could make use of the same two data sets gathered

by Hartshorn (2008) and the current researchers—in other words, the pretest and posttest essays written by McKay's and Hartshorn's students. These future researchers could divide the data by clauses or weighted clauses and recalculate complexity and accuracy statistics. The EFCR and WCR results could then be compared to the EFT/T results in Hartshorn's study and the current study.

It would be interesting to not only further test the validity of WCR as a construct for measuring accuracy, but also to see if perhaps using a different unit of analysis changes the complexity or accuracy results of both studies. It is possible that there would be greater statistical significance or larger effect sizes if a narrower unit of analysis than the T-unit were employed. For example, if a student's pretest essay could be divided into 40 T-units or 77 clauses, it is possible that division by clauses would result in a higher EFCR than the EFT/T ratio that would be produced by division into T-units. Thus, students' complexity and accuracy data may differ substantially according to which unit of analysis is used by researchers. A future study could look into this phenomenon.

Focus on manageability and practicality. The central *how* question regarding optimal accuracy instruction revolves around students: How can students be best helped or the error correction process made better for them? It would be interesting to investigate the *how* question from an instructor's perspective: How can dynamic WCF instructors be helped and the process made more practical or manageable for them, without hindering students' potential for improvement?

While McKay's students' accuracy showed no variation from Hartshorn's (2008) students' accuracy, it is possible that the variables of practicality or manageability, which were not quantitatively measured in either study, were affected. A numerical record of time spent by

both Hartshorn and McKay marking drafts outside of class would have been useful if such a record had been kept. Researchers could have compared time spent by each teacher marking drafts and determined whether or not McKay's method was truly more practical and manageable as he intended it to be, without negatively impacting students' accuracy.

So long as the accuracy of students experiencing dynamic WCF shows improvements, gathering data on teacher effort expended outside of class may contribute to the *how* of optimal error correction. For example, had numerical data on time spent outside of class marking drafts been kept by both Hartshorn (2008) and McKay, and McKay spent substantially less time than Hartshorn marking drafts but his students improved equally in accuracy, that would lend to an improvement of the practicality principle of dynamic WCF. Practitioners' experiences could possibly be enhanced, while the dynamic WCF process would still provide students with maximal benefits. A teacher who experiences burnout or lack of motivation resulting from too many drafts to mark will not be maximally effective in assisting students to improve their accuracy; thus, future researchers could take into account quantitative data regarding manageability to create a more practical model.

Conclusion

This study investigated a method of dynamic WCF dramatically modified from the way it was originally described by Hartshorn (2008). The results of this study suggest that teachers interested in improving their students' written grammatical accuracy should consider using dynamic WCF as an effective method for error correction. As long as teachers adhere to the core dynamic WCF principles of meaningful, timely, constant, and manageable, they can likely expect to see significant improvements in their students' linguistic accuracy with some flexibility in how the instructional methodology is carried out. This study also seeks to contribute

enlightening research regarding the *how* of effective error correction to currently-existing literature addressing the topic. The results of this study provide further evidence against Truscott's (1996, 2007) claims that error correction has no place in a classroom and give teachers increased confidence that improving students' linguistic accuracy is indeed possible.

References

American Council on the Teaching of Foreign Languages. (2012). *ACTFL Proficiency Guidelines 2012.* Alexandria, VA: ACTFL. Retrieved from http://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf

Bakeman, R., & Robinson, B. F. (2005). *Understanding statistics in the behavioral sciences.* Mahwah, NJ: Erlbaum.

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

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. doi: 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. doi: 10.1016/j.jslw.2005.08.001

Brown, S. & Larson-Hall, J. (2012). *Second language acquisition myths.* Ann Arbor, MI: The University of Michigan Press.

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

Celce-Murcia, M. (2013). An overview of language teaching methods and approaches. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 3-28). Boston, MA: National Geographic Learning/Cengage Learning.

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. doi: 10.1016/S1060-3743(03)00038-9

Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing, 13,* 345–348. doi: 10.1016/j.jslw.2004.05.001

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ellis, R., Sheen, Y., Murakam, 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. doi: 10.1016/j.system.2008.02.001

Evans, N. W., Hartshorn, K. J., Cox, T. L., & De Jel, T. M. (forthcoming). *Measuring written linguistic accuracy with weighted clause ratios: a question of validity.*

Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research, 14,* 1-19. doi: 10.1177/1362168810375367

Evans, N. W., Hartshorn, K. J., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System, 39,* 229-239. doi: 10.1016/j.system.2011.04.012

Farid, S. & Samad, A. A. (2012). Effects of different kind of direct feedback on students' writing. *Procedia – Social and Behavioral Sciences, 66,* 232-239. doi: 10.1016/j.sbspro.2012.11.265

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

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8,* 1-11. doi: http://dx.doi.org/10.1016/S1016-3743(99)80110-6

Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: where are we, and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing, 13,* 49-62. doi: 10.1016/j.jslw.2004.04.005

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge, England: Cambridge University Press.

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing, 22,* 307-329. doi: http://dx.doi.org/10.1016/j.jslw.2012.09.009

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

Hart, J. (2011). *Comparing the effects of two forms of dynamic corrective feedback on four characteristics of English language learner writing.* (Master's Thesis). Retrieved from http://contentdm.lib.byu.edu/cdm/singleitem/collection/ETD/id/2826/rec/1

Hartshorn, K. J. (2008). *The effects of manageable corrective feedback on ESL writing accuracy.* (Doctoral dissertation). Brigham Young University, Provo, UT. Retrieved from contentdm.lib.byu.edu/cdm/singleitem/collection/ETD/id/1526/rec/1

Hartshorn, K. J. & Evans, N. W. (2013, November). *Contextualizing the longitudinal effects of comprehensive corrective feedback on L2 writing accuracy.* Paper presented at the Second Language Research Forum, Provo, UT.

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly, 44,* 84-109. doi: 10.5054/tq.2010.213781

Hunt, K. W. (1965). *Grammatical structures written at three grade levels.* Urbana, IL: The National Council of Teachers of English.

Lee, S. (2009). *The efficacy of dynamic written corrective feedback on intermediate-high ESL learners' writing accuracy.* (Master's Thesis). Retrieved from http://contentdm.lib.byu.edu/cdm/singleitem/collection/ETD/id/2335/rec/1

McQuillan, J. (2012, January 20). *All correction, all the time: is written error correction worth the effort?* The Backseat Linguist. Retrieved from http://backseatlinguist.com/blog/?p=39

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing, 21,* 375-389. doi: http://dx.doi.org/10.1016/j.jslw.2012.09.004

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

Truscott, J. (1999). The case for "the case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing, 8,* 111-122. doi: http://dx.doi.org/10.1016/S1060-3743(99)80124-6

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16,* 255-272. doi: 10.1016/j.jslw.2007.06.003

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

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41,* 255-283. Retrieved from http://www.jstor.org/stable/40264353

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37,* 556-569. doi: 10.1016/j.system.2009.09.002

Shelley, A. (forthcoming). *Hit the ground running: Your guide to teaching linguistic accuracy.*

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

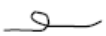
Wigglesworth, G. & Foster, P. (2008, April). *Measuring accuracy in second language performance.* Paper presented at the TESOL Convention, New York, NY.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* Manoa, HI: University of Hawaii Press.

Appendix A

Indirect Coding Symbols used by
Hartshorn and McKay

Hartshorn (2008) traditional method indirect coding symbols

D	= Determiner	S/PL	= Singular/Plural
SV	= Subject Verb Agreement	C/NC	= Count/Noncount
VF	= Verb Form	?	= Meaning is not clear
ro	= Run-on Sentence	AWK	= Awkward Wording
inc	= Incomplete sentence		= Word Order
VT	= Verb Tense	C	= Capitalization
PP	= Preposition	P	= Punctuation
SPG	= Spelling		= Omit
WF	= Word Form		= Something is missing
WC	= Word Choice	¶	= New Paragraph

McKay (2013) modified method indirect coding symbols

Fr-S	= Fragment—Missing Subject	VFg	= Verb Form—Gerund
Fr-V	= Fragment—Missing Verb	VFp	= Verb Form—Participle
∧	= Insert (Something Missing)	VFb	= Verb Form—Base
	= Omit	pp	= Preposition
spg	= Spelling	neg	= Negation
WF	= Word Form	s/pl	= Singular/Plural
WO	= Word Order	c/nc	= Count/Non-Count
WC	= Word Choice	PRO	= Pronoun Agreement
SV	= Subject Verb Agreement	det	= Determiner
VT	= Verb Tense	C	= Capitalization
aux	= Missing Auxiliary	“; . ! ?,”:=	Punctuation
VFi	= Verb Form—Infinitive		

Appendix B

Sample Prompts used by
Hartshorn and McKay

Hartshorn (2008) traditional method sample prompts

- Heroes
- Old and New
- Bridges
- Balance in Life
- Friendship
- Illness
- Wants and Needs
- Learning from Mistakes
- Perseverance
- Individualism
- Higher Education
- Work vs. Career
- Optimism
- Poverty vs. Wealth
- Respect
- Aging
- Democracy
- A Social Problem
- International Peace
- Humor
- Superstitions
- Effective Leadership
- Pranks or Double Meanings

McKay (2013) modified method sample prompts

- Many cultures have unique superstitions or myths. Describe a superstition or myths from your culture and how it affects people's actions or behaviors. Use specific details and examples to support your answer.
- Describe one difference in non-verbal communication methods between speakers of your native language and native English speakers. Why does this difference exist and how does this non-verbal communication pattern affect the culture? Use specific details and examples to support your answer.
- There is always a gap between the richest people in a country and the poorest. Do you believe this gap is increasing or decreasing in your country? Discuss the consequences of this change using specific details and examples to support your answer.
- Do you agree or disagree with this statement? “You can’t get rid of poverty by giving people money.” Use specific details and examples to support your answer.
- Microcredit refers to programs that give small loans to very poor people for self-employment projects. Most people who receive microcredit loans are women. How would you explain this fact? Use specific details and examples to support your answer.
- Do you agree or disagree with the following statement? *"University students should be required to give a few hours of community service each week."* Why or why not? Use specific details and examples to support your answer.

- Identify a product that is made in your country and exported to countries around the world. What makes this product a successful export good? What conditions in your country make it more efficient to produce this product than in other countries? Use specific details and examples to support your answer.
- Do you agree or disagree with this statement. *"The government should restrict cheaper imports from foreign countries to protect local business."* Use specific details and examples to support your answer.
- People write poetry as a way to express or release their emotions. What activity helps you to express or release positive or negative emotions? Use specific details and examples to support your answer.
- Poetry is an art form that has drastically changed in modern times. Consider another art form (music, dance, painting, theatre, sculpture) that is very different today than it was in the past. Use specific details and examples as you discuss what changes have occurred and why you think this art form changed?
- If you were an actor in a movie, would you prefer your character to be the hero (protagonist) or the villain (antagonist)? Use specific details and examples to explain your selection.
- Do you agree or disagree with this statement: *"Media from a country is a very accurate source of information about what life is really like in that country."* Use specific details and examples to support your answer.
- Describe either a successful or an unsuccessful effort you have seen a government make to conserve natural resources. Describe what the effort was, why the effort was a success/failure and how the effort could be more successful. Use specific details and examples to support your answer.
- Some people say that keeping animals in zoos is cruel; other people believe that zoos help protect animals. Which opinion do you agree with more? Use specific details and examples to support your answer.
- Do you agree or disagree with the following statement? *"One person cannot make a difference in improving the environment."* Use specific details and examples to support your answer.
- Some companies now make green products or products that are better for the environment. However, these products often cost more than similar non-green products. Would you pay more for a green product? Why or why not? Use specific details and examples to support your answer.

Appendix C

Error Tally Sheet used by Hartshorn (2008)

	TOO Much Freedom	Friendship	Solving Problems										Total
D	3	4	2										9
SV	1	1											2
VF	1	1	1										3
RO													
inc		1											1
VT	1	1											2
PP	3	4	3										10
SPG	3	2	3										8
WF	2	1	2										5
WC	3	1	1										5
S/PL	1	2	2										5
C/NC		1	1										2
?		1	1										2
AWK	1		1										2
WO	1												1
C													
P	1	2	3										6
omit		1	1										2
Insert	1	1	1										3
¶													
Score	7.5	7.8	7.4										

Appendix D

Error List used by Hartshorn (2008)

Error List

Determinates (D)

1. For example, it is unsafe when ***car*** drives too fast on urban roads.
2. Too much going on at ***a*** same time can cause some stress.
3. Actually, ***internet*** is being used by more and more people around the world.

Subject Verb Agreement (SV)

1. It always ***need*** to be for at least one hour.
2. It also ***increase*** the student's ability to learn.
3. My sunglasses ***was*** my most expensive purchase.

Verb Form (VF)

1. All of the assignments ***were been*** completed by the end of the day.
2. People should always be willing ***to working*** together.
3. You must believe in yourself so you do not ***would be failed***.

Appendix E

Edit Log used by Hartshorn (2008)

Ten-Minute Paragraph Edit Log

	Topics	Edits				
1	Too Much Freedom	→	→	→	✓	
2	Friendship	→	→	→	→	✓
3	Solving Problems	→	→	✓		
4	Lawyers	→	→			
5	Care for the Elderly	→				
6						
7						

Appendix F

Error Tally Sheet used by McKay (2013)

Cycle	1	2	3	1	2	3
Prompt	A	A	A	B	B	B
Error	#	#	#	#	#	#
Fr-S						
Fr-V						
^						
spg						
WF						
WO						
WC						
SV						
VT						
aux						
VFi						
VFg						
VFp						
VFb						
pp						
neg						
s/pl						
c/nc						
PRO						
det						
C						
“; . ! ?,”						
Error Total						

Appendix G

Error List used by McKay (2013)

Instructions: It is important that you are internalizing the feedback that you received responding to the prompts. This error log will help you to remember the feedback you received and process how it should change your expression in the future. List 10 inaccurate expressions; identify any error(s) present and write the expression with any error(s) eliminated. After you have completed the log, write a short reflection on the causes of these types of errors and how they can be avoided in the future.

#	Error(s)	Expression	
1		X:	
		O:	
2		X:	
		O:	
3		X:	
		O:	
4		X:	
		O:	
5		X:	
		O:	
6		X:	
		O:	
7		X:	
		O:	
8		X:	
		O:	
9		X:	
		O:	
10		X:	
		O:	Reflection:

Appendix H

Academic Input Inventory used by McKay (2013)

Instructions: Collect 10 academic phrases or structures that you could use in your writing on these topics. Adapt the key structures or phrases to your own academic sentence. It will be very useful if you have a native speaker or more advanced learner check the accuracy of your sentences.

1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	