https://doi.org/10.32038/frsl.2023.01.02

Dynamic written corrective feedback frequency and its effects on ESL writing fluency, accuracy, and complexity[1]

K. James Hartshorn*
Brigham Young University, USA
Suzanne H. Rice
Brigham Young University, USA
Grant Eckstein
Brigham Young University, USA
Norman W. Evans
Brigham Young University, USA

Correspondence
Email: james_hartshorn@byu.edu

Abstract

Dynamic written corrective feedback (DWCF) is an instructional strategy that uses error codes to address patterned errors in student writing. While studies have shown that DWCF significantly increases the accuracy of student writing, less is known about the effect frequency of teacher feedback has on accuracy, fluency, and complexity. This study observed two treatment groups who wrote short compositions and received comprehensive feedback daily (n=34) or every-other-day (n=34) with the amount of writing held constant at twenty minutes per week. When compared with a control group (n=33), findings showed greater accuracy for both the every-other-day group (p<.001, d=1.07) and the daily group (p<.001, d=1.26) while the daily group had significantly larger gains in fluency compared to the control group (p<.001, d=1.15) and the every-other-day group (p<.001, d=1.14). Two forms of complexity were also examined. No statistically significant differences were observed for clauses per T-unit. For mean length of T-unit, there was no significant difference between the control group and the daily group, but there was a significant decline for the every-other-day group compared to the control group (p=.025, d=0.75). Results contribute to SLA theory and can be useful for practitioners seeking to improve student writing in regular, short writing tasks.

ARTICLE HISTORY

Received: 17 February 2023
Revised: 28 September 2023
Accepted: 27 October 2023

KEYWORDS

dynamic written corrective feedback, frequency, fluency, accuracy, complexity

[1] An earlier, unpublished draft of this work can be found at: https://scholarsarchive.byu.edu/etd/9454

Introduction

Feedback on student writing is considered a critical component of writing development (Hyland & Hyland, 2019). Many students demand it, and many writing teachers spend considerable time responding to it (Polio, 2010) even when they view it as a drudgery (Lee, 2011). While there is conflicting evidence that feedback leads to actual improvement in student writing (Biber et al, 2011; Ferris, 1997; Hyland & Hyland, 2019), less conflicting evidence is associated with the efficacy of written corrective feedback (WCF), or feedback designed to improve writing accuracy (Bitchener & Knoch, 2010a, 2010b; Kang & Han, 2015; Lim & Renandya, 2020). This is despite Truscott's (1996) claim that WCF may be harmful and should be abandoned altogether (see also Truscott 1999, 2001, 2007). In reaction to Truscott, writing scholars and practitioners have created study after study demonstrating the benefits of WCF. So much has been published that researchers have turned to meta-analyses to aggregate findings (see Biber et al, 2011; Lim & Renanyda, 2020; Kang & Han, 2015). The general consensus of the field (excepting Truscott's unpublished paroxysms, e.g., Truscott, 2020, 2022) is that WCF leads to gains in student learning and performance in many situations.

In the last few decades, researchers have shifted their attention to exploring the best situations and formats for providing WCF, thanks largely to Ferris' 2004 call to do so. They have done this by questioning whether direct or indirect feedback may be best (Bitchener et al., 2005; Bitchener, 2021; Ferris & Roberts, 2001; Van Beuningen, 2008, 2010) and whether comprehensive or focused feedback may be better (Bitchener & Knock, 2008; Sheen et al., 2009; Van Beuningen et al., 2012). Some studies have examined timing of feedback, questioning whether feedback should be given immediately after performance or whether it can be delayed (Arroyo & Yilmaz, 2018; Eckstein et al., 2020; Shintani & Aubrey, 2016).

Despite the volumes of research, there has been little investigation of the effect of feedback frequency on student performance (Lavolette et al., 2015; Quinn, 2021). Thus, the question of how frequently a student should receive WCF—should it be daily, every other day, or less frequently—remains underexplored in WCF studies. While cognitive scientists have explored spacing effects and found benefits of distributed practice (see Cepeda et al., 2006; Rohrer & Pashler, 2007), and language researchers have demonstrated benefits of massed practice episodes (see Freed et al., 2004; Serrano and Muñoz, 2007; Spada & Lightbown, 1989), these studies largely overlook grammar feedback and are generally concerned more with skill practice. Simply using a more-is-better approach to feedback that ignores feedback frequency may overwhelm students or waste valuable time. Furthermore, if feedback frequency contributes to language gains, then it can potentially inform theories of language acquisition.

The present study explores how WCF frequency may impact writing development, particularly in relation to written complexity, accuracy, and fluency. Our aim is to

evaluate whether daily DWCF (dynamic written corrective feedback) is more beneficial than DWCF provided every other day over the course of a semester. Ultimately, results can inform both language acquisition theory and classroom pedagogy as teachers seek the best ways to use their limited instructional time.

Literature Review

A defining feature of most second language writing instruction is the use of error correction. According to Evans et al. (2010), 92% of over 1,000 ESL and EFL teachers surveyed specified that they used some kind of WCF in their writing classes. They reported that, on average, 44% of the teachers' time was used to provide accuracy-based feedback. Lee (2003) found similar evidence of teachers pouring time into error correction. According to her study of more than 200 writing teachers in China, 60% agreed that it is the teacher's responsibility to mark errors, and in follow-up interviews, the most frequent comment was that the practice was time consuming. Indeed, 61% of survey respondents reported spending more than 20 minutes marking each advanced composition. In a related study, Lee (2009) found that despite teacher claims that feedback should address content and form, in reality, of the 5,353 points of teacher feedback analyzed, form feedback (94.1% of all marks) almost completely eclipsed feedback on content (3.8%) and organization (0.4%).

The extraordinary effort used to correct errors has led researchers to question the best ways to provide it. Direct feedback, in which teachers give the correct form for each error, provides students with unambiguous error correction (Bitchener 2008; Ellis et al., 2008; Nassaji, 2015) while indirect feedback, where a code, an underline, or some other mark is used to alert a writer of an error, requires the student to reflect on linguistic forms to address the error rather than simply accept the teacher's correction uncritically (Ferris, 2001; 2003; Lelande, 1982). While studies have shown that both approaches may be effective (see Kang & Han, 2015; Karim & Nassaji, 2019), many researchers agree that the effectiveness of one type over the other is impacted by factors such as language proficiency, the "treatability" of the error, the instructional context, the intensity of the feedback, and other individual differences (Bitchener, 2021; Bitchener & Knoch, 2009; Ferris & Roberts, 2001; Kang & Han, 2015; Karim & Nassaji, 2019; Liu & Feng, 2023; Storch, 2021; Utsi, 2023).

Teachers also need to determine whether to correct a small set of errors or correct all errors—this is often described as comprehensive or unfocused feedback (Sheen et al., 2009). While findings from studies of comprehensive feedback have been mixed (see Bitchener, 2021), some scholars have favored focused feedback as a means of preventing attentional overload. Teachers have also been concerned with the emotional toll error correction may have on students (Han & Hyland, 2019). Furthermore, focused feedback tends to improve accuracy of the specific structures targeted for feedback (see Bitchener, 2021).

Nevertheless, many scholars have criticized focused feedback because it lacks ecological validity for classrooms where students expect to overcome many different error types (e.g., Ferris, 2010; Hartshorn et al., 2010; Hartshorn & Evans, 2015; Nassaji, 2016; van Beuningen, 2021, van Beuningen et al., 2012; Xu, 2009). Xu (2009) has cautioned that a narrow emphasis on focused feedback that may improve accuracy in one limited area could also have unintended, detrimental consequences in other areas. Similarly, van Beuningen (2021) has warned that relying mainly on focused feedback could "mislead learners about the amount of effort and attention it takes to monitor and self-edit their language production" (p. 307). Research in this area continues though methodological concerns persist. For example, while Craven (2023) found no meaningful effects of unfocused feedback on fluency, accuracy, or complexity in new texts following treatment, the study only included four rounds of feedback across 14 weeks, and the pieces of writing were likely too large for students to effectively process the high volume of feedback provided.

Teachers must also wrestle with the practicality of providing WCF in terms of timing. While it is generally assumed that immediate feedback is always best, feedback cannot always be instantaneous, nor is this always preferrable. In a metanalysis of 53 verbal learning experiments investigating the timing of feedback, Kulik and Kulik (1988) noted that half of the studies reported a benefit to delayed feedback, and the other half did not. In contrast, the study of written corrective feedback timing is a recent development (Quinn, 2021). Eckstein et al. (2020) investigated timing of DWCF feedback among graduate ESL writers by dividing students into a *timely* group who received immediate WCF feedback every other day for 14 weeks and a *postponed* group who received feedback on all of their weekly writing during the last two weeks of the semester. Results showed that neither group outperformed the other in measures of writing accuracy, though timely feedback did produce writing that was more fluent and complex.

Using CALL software in a semester-long study, Lavolette et al. (2015) provided WCF for four essays in the form of highlighting and metalinguistic information to an immediate feedback group within forty minutes of submitting their writing, while a delayed feedback group received their feedback between one to three weeks later. Using error-free T-units as their metric of accuracy, they found no differences between the two groups. Using Google Docs to contrast the effect of synchronous versus asynchronous WCF, Shintani and Aubrey (2016) provided focused feedback regarding the hypothetical conditional for two writing tasks to synchronous and asynchronous experimental groups while a comparison group received no feedback. Findings from a pretest, immediate posttest, and a delayed posttest showed that while both experimental groups improved their accuracy of the target structure, only the synchronous group outperformed the comparison group.

Similarly, Arroyo and Yilmaz (2018) also examined the efficacy of synchronous and asynchronous WCF using direct, focused feedback on Spanish noun-adjective gender

agreement. A synchronous group received immediate feedback in less than a minute via text, the asynchronous group received feedback after the completion of the task, and the control group received no feedback. Though both the synchronous and asynchronous groups outperformed the control group, there was no meaningful difference between treatment groups on a grammaticality judgement test. Though it appears that in some cases feedback timing may be less critical so long as feedback is provided, more study is needed to better understand the effects of feedback timing.

Dynamic Written Corrective Feedback

One approach developed to provide indirect, comprehensive, and timely feedback is that of Dynamic Written Corrective Feedback (DWCF; see Hartshorn et al, 2010). Based on skill acquisition theory, DWCF is an instructional strategy designed to address many of the limitations associated with traditional approaches to feedback. DWCF was intended to help students improve the accuracy of their writing without the large processing demands that might undermine learning or overwhelm students emotionally. Though feedback is comprehensive, a crucial mechanism is constant (i.e., daily) feedback limited to a small amount of student writing on each occasion. Smaller pieces of writing completed each class day limit teacher feedback while providing students with manageable amounts of input. The expectation is that writing accuracy will improve over time through many iterations of student writing, immediate teacher feedback, and additional opportunities to apply ongoing feedback. Because the feedback is based on short pieces of writing, the cycle is manageable for both teacher and student, making it possible to repeat the process on a daily or on a near daily basis.

Skill acquisition theory suggests that as frequent instruction, practice, and feedback increase, errors are likely to decrease. Student writers will move from declarative knowledge to procedural knowledge to automatization (i.e., fluent and accurate writing) (e.g., DeKeyser, 2007; Hartshorn & Evans, 2015). Typically, the manageability of this process is ensured by limiting student writing to ten minutes each class period, having teachers provide comprehensive, indirect feedback via error codes, and then having students continue to edit their work based on ongoing teacher feedback until it becomes free of errors. In the meantime, students keep tallies of the specific error types they produce, and the aggregation of this information informs and prioritizes classroom instruction and activities provided by the teacher (Evans et al., 2010; Hartshorn et al., 2010; Hartshorn & Evans, 2015).

Efficacy of DWCF

Greater written accuracy from DWCF has been seen in studies of pre-matriculated ESL writers (e.g., Evans et al., 2010; Hartshorn & Evans, 2012, 2015; Hartshorn, et al., 2010) as well as studies of university-matriculated ESL writers (e.g., Evans et al., 2011; Kurzer, 2018). Hartshorn et al. (2010) observed that pre-matriculated students significantly improved in error-free T-units over 15 weeks of daily DWCF intervention compared to a control group, and Evans et al. (2010) similarly noted a significant improvement in error-free T-units in 13 weeks of DWCF treatment. Evans at al. (2011)

and Hartshorn and Evans (2015) conducted another 15-week and 30-week study respectively with undergraduate ESL writers and found DWCF to be effective in reducing students' overall errors among these learners as well. Hartshorn and Evans (2012) observed that DWCF led to improvements with determiners, mechanics, and semantic accuracy. In a 10-week study of university writers, Kurzer (2018) found that DWCF led to improvements in mechanical, local, and global errors.

Evidence for the efficacy of DWCF has also been observed in EFL contexts. For example, in Arabic-speaking nations, Mahmoud (2023) found positive student sentiment toward DWCF because of the way in which it directed their attention toward critical thinking to help them learn to correct their own errors. Similarly, Khalifa (2022) observed higher writing scores for those receiving DWCF compared to a control group. Moreover, for native speakers of Persian, Kamalian and Lashkarian (2014) noted that their treatment group receiving DWCF outperformed their control group in terms of writing accuracy. Furthermore, though Marzban and Arabahmadi (2013) observed improved accuracy with minimal effects on fluency and complexity, Sayad Deghatker et al. (2022, 2023) observed increases in accuracy, fluency, and complexity with DWCF.

Additional studies have shown improved accuracy in languages other than English as a second language. For example, in a Korean as a foreign language context, Oh et al. (2020) and Cho (2021) observed improvements in accuracy for DWCF groups compared to a control group, though An (2019) saw no differences in accuracy between a direct feedback group and a DWCF group. It should be noted, however, that in this study there were only nine writing-feedback exchanges, many fewer than the forty to fifty exchanges more typical for DWCF research. In the Japanese as a foreign language context, Akiyama and Fleshler (2013) tested first-year students representing a much lower proficiency than typically examined in DWCF research. These scholars found that while DWCF did not improve the accuracy of all linguistic features, students wrote more accurately than those in the control group in terms of particle use and the formation of predicates. Qualitative data suggests that these student experiences with DWCF were largely positive.

Feedback Frequency

While research has shown the efficacy of DWCF, some practitioners have continued to question its practicality given that daily writing and feedback is integral to the approach. McQuillan (2012) questioned the rationale of what he called an "all-correction, all-the-time agenda" (para. 32), and Casanave (2007) criticized it as "impractical," particularly for classes that meet only once a week or have large enrollments. Teachers like Shelly (2014) and Eddington (2014) have similarly acknowledged challenges providing frequent feedback using DWCF even with small classes sizes when the teachers have limited teaching experience.

Accordingly, one approach to addressing manageability could be to reduce the frequency of writing-feedback episodes, yet little is known about how this might affect error correction. Certainly, cognitive scientists have been interested in practice frequency and have explored the effects of massed and distributed practice episodes. Massed practice, in which episodes occur close together (such as daily practice), has been contrasted with distributed practice that might occur every few days or weekly. The results have been mixed. Some studies in applied psychology across a range of educational domains favor distributed practice (i.e., Cepeda et al., 2006; Rohrer & Pashler, 2007) while other studies of general language teaching have demonstrated greater learning with massed practice experiences (Netten & Germain, 2004).

Spada and Lightbown (1989) compared ESL learners in a five-month intensive program with comparison groups who received a comparable amount of English instruction over several years. The intensive group outperformed the comparison groups on measures of listening and reading comprehension as well as oral fluency. In a follow-up study in which ESL instruction was massed into five months of instruction or distributed across ten months, the massed group again outperformed the distributed group on vocabulary, listening, and composition tests (Collins et al., 1999). Freed et al. (2004) similarly found that a seven-week massed French immersion group outperformed a twelve-week distributed learning group on oral fluency, and Serrano and Muñoz (2007) examined EFL learners who studied for 110 hours in increments of seven months, three to four months, or just two months and found that "concentrating the hours of English instruction in shorter periods of time is more beneficial for...students' learning than distributing them in many months" (p. 305).

In the only DWCF study of feedback frequency, Kurzer (2018) found that students improved their writing accuracy whether they wrote and received feedback approximately twice per week (massed), about once a week, or every other week (distributed). In this case, comprehensive grammar practice improved efficacy regardless of feedback frequency. Kurzer noted that the findings of his study suggested that "DWCF may not need to be quite as frequent as originally proposed" (p. 25). In follow-up surveys, Kurzer observed that teachers believed that daily writing and revision was "more often than needed" and that teachers "adapted the frequency" for their context (Kurzer, 2018, p. 13).

In other research, Lam et al. (2011) studied four different feedback groups based on 2, 4, 7 and 14 feedback sessions during a 70-minute exercise. The group with the most feedback became overwhelmed and demonstrated the lowest performance. In contrast, the group receiving the moderate-high frequency of feedback outperformed the other groups. These findings suggest that there may be an "optimum" frequency for feedback (Lam et al., 2011, p. 226). Accordingly, this current study seeks to identify the optimum level of feedback frequency needed to optimize the appropriate development of written accuracy, fluency, and complexity. Scholars have encouraged more research in these areas (Quinn, 2021; Mao & Lee, 2020), and Lee (2019) has

emphasized the need for research of the factors that influence the efficacy of written corrective feedback. We believe that frequency may be one factor that could influence the effectiveness of DWCF.

Complexity, Accuracy, and Fluency

Although accuracy is the primary focus of the current study, accuracy should be viewed within the broader context of fluency and complexity. The trade-off hypothesis suggests that limited attentional capacity and working memory allow for cognitive competition between complexity and accuracy on the one hand and fluency on the other. Tensions can also occur between complexity and accuracy (Skehan, 2009). The hypothesis predicts that typically learners are likely to observe increases in accuracy and fluency, or complexity and fluency, but not both (p. 512). However, the cognition hypothesis rejects the notion of a limited attentional capacity. Robinson (2003) describes learners who are pushed by challenging tasks to develop greater complexity, improved accuracy, and an enhanced ability in memory and attention. The demands of such tasks may produce an opportunity for learners to access multiple attentional pools simultaneously (Robinson, 2003).

Housen et al. (2012) have described complexity, accuracy, and fluency as "distinct and competing areas of L2 performance" (p. 3). They note that a learner's L2 productive complexity is an indication of how students internalize new language elements. The degree of their linguistic accuracy shows the refinement of their interlanguage towards more target-like productions. The proceduralization and consolidation of language knowledge is evident in a learner's fluency.

Within the process of DWCF, as writing becomes automatized and more accurate, more vocabulary and complex structures can be produced more quickly. Accuracy and complexity are an indication of language knowledge while fluency is considered a learner's control over that linguistic knowledge (Housen & Kuiken, 2009). The trade-off hypothesis predicts that when the focus is on fluency, it competes with accuracy, which then in turn competes with complexity (Skehan, 2009). This is consistent with the findings of many DWCF studies (Eckstein et al, 2020; Evans et al., 2010; Evans et al., 2011; Hartshorn & Evans, 2012; Kurzer 2018).

Based on this review, the present study examines two reduced forms of DWCF compared to a control group. One treatment group engaged in five minutes of writing four days per week (massed) and another group engaged in 10 minutes of writing two days per week (distributed), both receiving immediate, comprehensive, indirect feedback on 20 minutes of writing weekly. The study was motivated by the following research questions:

RQ_1: What is the effect of the reduced forms of DWCF in terms of accuracy, fluency and complexity of ESL student writing compared to a control group not receiving DWCF?

RQ2: How comparable are the two reduced forms of DWCF with different frequencies in terms of accuracy, fluency, and complexity of ESL student writing?"

Methods

The current study was intended for six intact grammar classes within an intensive English program. Three classes were at the intermediate mid to intermediate high proficiency level and three classes were at the intermediate high to advanced low proficiency (American Council on the Teaching of Foreign Languages, 2012). The control group was made up of one class from each of the proficiency levels as were the two treatment groups. In total, 101 students participated in the study. In addition to the grammar class, all participants also had a separate writing class, which included larger writing assignments along with nominal feedback in areas such as organization, rhetorical structure, and content, as well as some minimal feedback areas such as word choice, grammar, and mechanics.

Participants

The control group (*n*=33) included 17 males and 16 females, and ages ranged from 18 to 32 (*M*=24). The L1 backgrounds of students in the control group included Spanish (18), Portuguese (6), Japanese (4), Chinese (2), Korean (1), Central Khmer (1), and Bambara (1).

The first treatment group (*n*=34) (we refer to as the "every-other-day" group) had 16 males and 18 females with their ages ranging from 18 to 42 (*M*=25). Similar to the control group, they were primarily native speakers of Spanish, with Portuguese and Japanese speakers in similar ratios including Spanish (21), Portuguese (4), Japanese (4), Chinese (2), Arabic (2), and Korean (1).

The second treatment group (*n*=34) (we refer to as the "daily" group) was made up of 19 males and 15 females with ages ranging from 19 to 38 (*M*=24). Student L1 backgrounds included Spanish (21), Portuguese (5), Japanese (2), Chinese (2), Creole (2), Korean (1), and Russian (1). Each of the three groups were fairly similar in terms of age, the proportion of male and female participants, and number of speakers from different L1 backgrounds.

Method of Instruction

The control group was provided traditional grammar instruction using a common textbook in accordance with the IEP curriculum. Feedback provided to the control group was largely based on answers to textbook exercises rather than on student written production. In their writing classes, students in the control group wrote several compositions during the semester using common process writing approaches. They also wrote three thirty-minute essays throughout the semester but were only given feedback on organization and content for those writing tasks.

The every-other-day group received one version of DWCF. These students wrote for ten minutes on a given topic on Monday. Then on Tuesday, they received feedback on their writing in the form error codes. On Wednesday, they made corrections and then wrote for another ten minutes on the same topic. On Thursday, the writing was returned with additional error codes if needed, and the final draft was due on Monday. Participants also tracked their error types in an effort to raise their awareness of the linguistic features that were most problematic for them. Students worked with one distinct prompt each week. In total, participants wrote for twenty minutes in class each week and responded to feedback twice.

Students in the daily group used a slightly different version DWCF. They only wrote for five minutes on Monday. They received feedback from their teacher before the start of Tuesday's class, made appropriate corrections in class, and then wrote for an additional five minutes on the same theme. This process was repeated on Wednesday and Thursday. The final version of the paragraph was due on Monday. Students in the daily group also tracked error types. Like the other treatment group, the daily group wrote for a total of twenty minutes in class, but they responded to teacher feedback four times during the week.

Class instruction was not based on a linear, chapter by chapter progression through the textbook. Rather, instruction for the treatment groups was prioritized based on what students needed most as evidenced by the kind of feedback provided. Students in the treatment groups typed their pieces of writing and made corrections on computers. They used proprietary software which enabled the use of cut, paste, and delete functions. It did not allow access to software that could check their spelling or grammar. Table 1 compares the regimens of the two treatment groups. It is important to emphasize that though the frequency of the feedback differed between the two treatment groups, the actual writing time (and the subsequent volume of feedback) each week was the same for both groups.

Table 1

Comparison of the 3 Writing Groups with the Original DWCF Process

	Study Groups		
Group differences	Control	Every-other-day	Daily
Daily writing minutes	0	10	5
Writing days per week	0	2	4
Feedback times per week	0	2	4
Weekly writing minutes	0	20	20

Pre-tests and Post-tests

Before and after the participants engaged with the 14-week DWCF treatment, they completed pre-tests and post-test essays. The pretest was made up of one thirty-minute essay and four ten-minute paragraphs for a total of 70 minutes of writing on five different topics. Students produced the same amount of writing for the post-test

on five different topics. This aspect of the study is important since it strengthens the research in several ways. Different prompts might produce different results depending on student familiarity with specific topics (Lee & Anderson, 2007). Thus, with multiple prompts and writing occasions, poor performance on either the pretest or posttest is less likely to be due to a student's unfamiliarity with a given topic. Table 2 presents the writing prompts for the pretest and posttest. Most topics remained within the personal sphere of interest, and participants were also asked to perform various language functions such as description, narration, and support of an opinion that, according to the ACTFL (2012) proficiency scale, conform with what L2 learners should be able to do at the intermediate high to advanced low levels. Although the results of many WCF studies have been based on small amounts of writing, the results of this study should be robust in terms of reliability due to the sheer volume of data rated and recorded across varied topics and genres (Evans et al., 2014).

Table 2

Prompt Topics

	Pretest	Posttest
30 Minute	Describe the most important animal to your country and explain why.	Agree/Disagree Colleges should require students to study abroad.
10 Minute 1	Describe a place you would go to relax and explain why.	Describe a person that helps you to be happy and explain why.
10 Minute 2	Agree/Disagree Only people who earn a lot of money are successful.	Describe an important character trait that is needed to be successful in life.
10 Minute 3	Describe a meaningful experience before you came to this school.	Describe a difficult experience and what you learned.
10 Minute 4	Agree/Disagree Class attendance should be mandatory.	Agree/Disagree Teachers should assign homework every day.

For optimal comparability, this study examined a variety of language measurements consistent with other DWCF studies (Evans et al., 2010; Evans, et al., 2011; Hartshorn, et at., 2010; Hartshorn & Evans, 2015). Accuracy was defined as the ratio of error-free clauses to the total number of clauses (Wigglesworth, 2008). This measure was chosen because of its appropriateness for advanced learners (Evans et al., 2014), and it is used in many previous DWCF studies. Fluency was defined as the total number of words written within a set time period (Wolfe-Quintero et al., 1998). Complexity was defined in two ways. Each used the T-unit, developed by Hunt (1965), which he defined as "one main clause plus the subordinate clauses attached to or embedded within it" (p.14). The first measure of complexity was the mean length of T-unit (MLTU), and the second measure of complexity was the number of clauses per T-unit (Wolfe-Quintero, et al., 1998).

Two individuals rated 65% of the pretests and posttests with 30% overlap. The writing samples they both worked on were randomly taken equally from the high, middle, and low performing students' writing based on Rasch modeling fair averages from the pretest. Raters followed four basic guidelines when determining what constituted a new T-unit. They were that punctuation could be ignored in favor of meaning, transition words like *however* and *nevertheless* could begin a new T-unit, a change of topic signaled a new T-unit, and subordinating clauses must clearly attach to the main clause to be included. After reliability estimates were established, where differences were observed across raters, means were calculated and used in the subsequent analyses. Table 3 shows the interrater reliability (Pearson correlation coefficients between the two raters) for each measure on the pretest and posttest.

Table 3

Inter-rater Reliability for Accuracy and Complexity Measures

Data Measure	Pre-test	Post-test
Error-free Clause Ratio	.924	.896
Words per T-unit	.981	.969
Clauses per T-unit	.979	.955

To determine the relative equivalence of the groups prior to further statistical analysis, pretest scores were compared using one-way ANOVA. Findings revealed no differences for accuracy, $F(2,98)=1.620$, $p=.203$, fluency, $F(2,98)=.278$, $p=.760$, or complexity as measured by mean length of T-unit, $F(2,98)=1.959$, $p=.146$. However, a moderate difference was observed for complexity as measured by clauses per T-unit, $F(2,98)=3.77$, $p=.026$. A Tukey post hoc test revealed that this difference was based on slightly more clauses per T-unit produced by the every-other-day group ($M=2.08$, $SD=.378$) in contrast to the control group ($M=1.90$, $SD=.234$), $p=.04$, $d=0.61$. The reason for this unexpected difference is not clear. However, we decided to proceed with the planned analyses since most of these measures suggested the comparability of these groups and no differences were observed across groups in terms of accuracy—the main focus of this study. Moreover, we expected that the analyses of variance would account for any minor pretest differences across groups.

Results

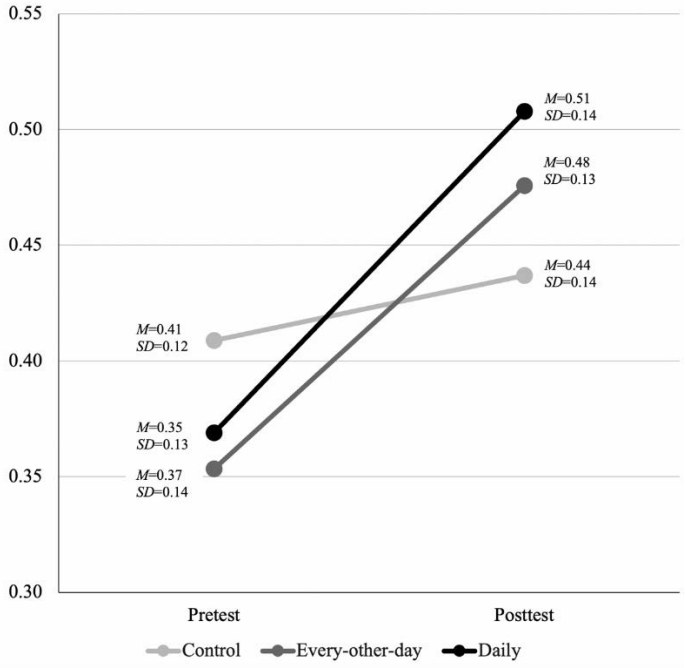
In order to answer the research questions, one-way ANOVA was used to compare gains scores across the three groups for the four language measures examined in this study.

Accuracy

Accuracy as defined by error free clause ratios varied across groups, $F(2,98)=14.103$, $p<.001$. Results of a Tukey post hoc test showed that the control group experienced a significantly smaller improvement in accuracy ($M=0.028$, $SD=0.075$) compared to the every-other-day group ($M=0.122$, $SD=0.099$), $p<.001$, $d=1.07$, and the daily group ($M=0.139$, $SD=0.099$), $p<.001$, $d=1.26$. Figure 1 illustrates pretest and posttest plots for accuracy.

Figure 1

Pretest and Posttest Accuracy by Group

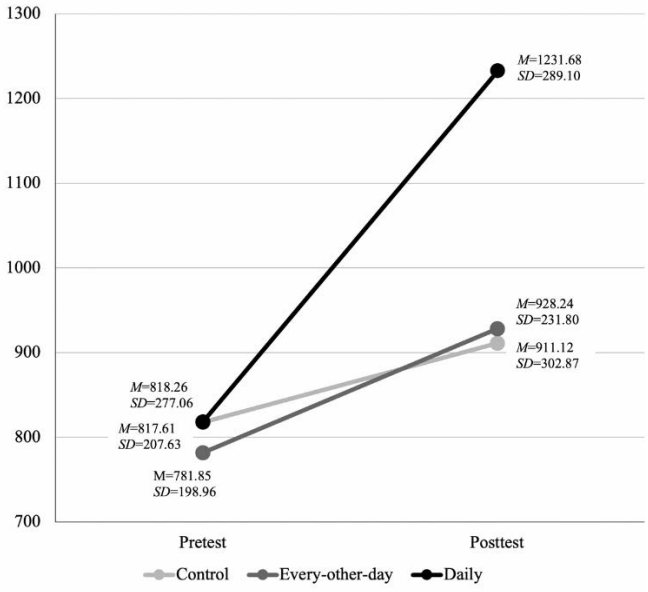


Fluency

Fluency as defined by the total number of words produced also varied across groups, $F(2, 98)=15.336$, $p<.001$. A Tukey post hoc test indicated that the daily group experienced significantly larger increases in fluency ($M=413.412$, $SD=264.322$) compared to the control group ($M=93.515$, $SD=292.727$), $p<.001$, $d=1.15$, and the every-other-day group ($M=146.382$, $SD=197.82$) $p<.001$, $d=1.14$. Figure 2 shows pretest and posttest plots for fluency.

Figure 2

Pretest and Posttest Fluency by Group

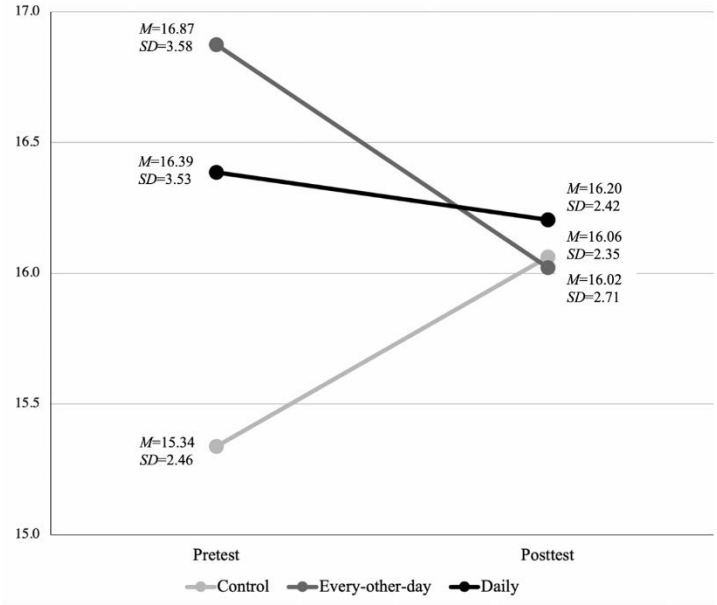


Complexity I

Complexity as defined by the mean length of T-unit differed across groups, $F(2,98)=3.545$, $p=.033$. A Tukey post hoc test showed the difference as being between the gains in the control group ($M=0.725$, $SD=1.584$) and the loss in the every-other-day group ($M=-0.852$, $SD=2.509$) $p=.025$, $d=0.75$, though there was no meaningful difference between the control group and the daily group ($M=-0.852$, $SD=2.509$), $p=.283$. Figure 3 illustrates pretest and posttest plots for the MLTU.

Figure 3

Pretest and Posttest MLTU by Group

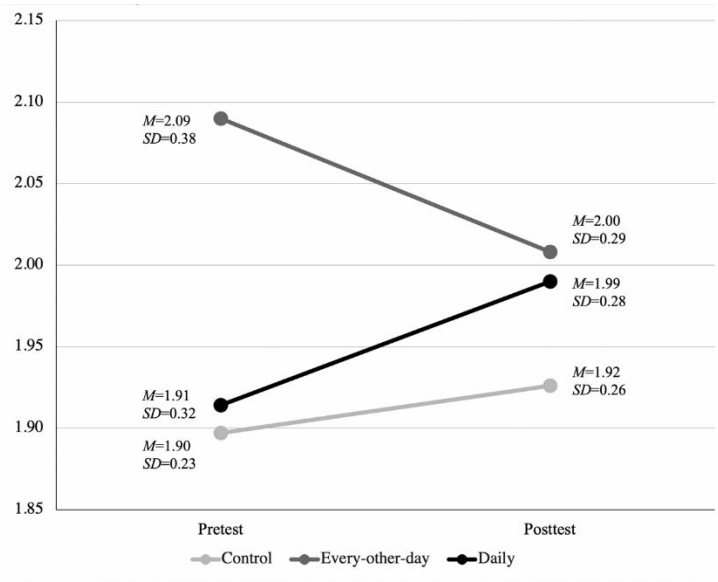


Complexity II

Complexity as defined by the number of clauses per T-unit was not statistically significant across groups based on a significance level of .05, $F(2,98)=2.956$, $p=.057$. A Tukey post hoc test indicated a moderate effect size between the every-other-day group ($M=-0.081$, $SD=0.32$) and the daily group ($M=0.077$, $SD=0.307$), $p=.051$, $d=0.5$. No difference was found between the control group and the every-other-day group or between the control group and the daily group. Figure 4 shows the pretest and posttest plots for the clause per T-unit ratio.

Figure 4

Pretest and Posttest Clause per T-Unit Ratio by Group



Discussion

The intent of this study was to examine the effects of the frequency of teacher feedback on L2 writing accuracy, complexity, and fluency. The first research question investigated whether a control group who received no accuracy-based feedback performed differently from those with DWCF interventions in terms of accuracy, fluency, and complexity. Results showed that while the control group experienced a slight improvement in accuracy over a semester, the gains were nearly negligible and significantly smaller than accuracy gains of the two DWCF groups examined in this study. This corroborates previous DWCF research which similarly showed DWCF groups outperforming control groups regarding accuracy (Hartshorn et al. 2010; Hartshorn & Evans, 2012, 2015; Kurzer, 2018).

In terms of fluency, the daily group significantly outperformed the control group, but there was no difference between the control and every-other-day group. This combination suggests that fluency in writing developed more when learning was massed on a daily basis rather than distributed with writing breaks between practice episodes, echoing previous research that favors massed practice (e.g., Freed, 2004; Netten & Germain, 2004; Serrano & Muñoz, 2007). Moreover, the results corroborate Eckstein et al.'s (2020) study in which "timely" feedback episodes resulted in greater fluency gains than postponed feedback, presumably because more regular practice with accountability correlates with greater idea development and writing habits, an effect observed in studies of professional academic writers (e.g., Boice, 1984; Gray & Birch, 2000).

Complexity findings were somewhat mixed in that the two complexity measures revealed slightly different patterns. No significant difference was found between control and treatment groups in terms of clauses per T-unit measures, indicating that

DWCF intervention did not affect clause production; nor did it, for that matter, reduce complexity. Thus, while accuracy increased, complexity, as measured as clauses per T-unit, was not affected as has been reported in previous studies (Hartshorn et al., 2010; Hartshorn & Evans, 2015). On the other hand, the control group demonstrated greater development in complexity as measured by MLTU compared to the every-other-day condition while there was no significant difference found between control and daily writing condition. This suggests that daily writing may have encouraged shorter T-unit production while no grammar intervention may have led to longer T-unit production. This complexity result supports the trade-off hypothesis (Skehan, 2009), at least when DWCF is provided in massed practice situations.

In sum, results show that DWCF interventions lead to accuracy, fluency, and complexity gains relative to a control group. However, the type of DWCF intervention (i.e., massed verses distributed) appears to have an impact on whether fluency and complexity increase, which is discussed in more detail below.

The second research question explored the comparability of the two DWCF treatments. Findings show that the daily (or massed) group who consistently wrote for 5 minutes in each class, demonstrated the highest gains in writing accuracy and fluency. This may be important for L2 practitioners if learners increase in their written accuracy with less time allocated for the DWCF process in class and less time needed for teachers to mark student writing. Students would also have less feedback to process each day. Lee (2019) observed,

> "*With fewer underlines, circles, error codes, etc., WCF is less confusing and intimidating, and students are more likely to take risks (e.g. experimenting with new language to express meaning rather than use grammatically safe expressions that do not necessarily convey their real intended meaning) and as a result build fluency and confidence in writing." (p. 527)*

These results correspond with previous DWCF studies that reported significant gains in accuracy (Evans, Hartshorn, McCollum & Wolfersberger, 2010; Evans, et al., 2011; Hartshorn, et al. 2010, Hartshorn & Evans, 2012, 2015; Kurzer, 2018), particularly at the intermediate mid to advanced low proficiency. One benefit of this study, then, may be the finding that accuracy can still be improved in half the time used in the original DWCF process. Skill acquisition theory implies that the more a skill is practiced the more proficient a learner can become. Both treatment groups wrote for the same amount of time. How the time was divided made little difference in terms of accuracy. Students in both treatment groups significantly improved the accuracy of their writing. These findings support those of Kurzer (2018) who found that all massed and distributed practice conditions led to improved student writing.

The schedule of writing and feedback every other day (i.e., distributed practice) also seems to increase the accuracy of student writing. This is another useful finding for

practitioners, particularly for classes that do not meet every day. Nevertheless, these findings showed that fluency did not increase for the every-other-day group in contrast to the daily group. Since both treatment groups experienced the same amount of writing time and feedback, perhaps the difference in number of writing and revision occasions put the every-other-day group at a slight disadvantage. Writing and receiving feedback four times a week, even for just five minutes, seemed to make more of a difference than writing twice a week for ten minutes. This suggests the advantage of massed practice over distributed practice in terms of fluency development. Teachers will need to determine what schedule may work best for them and may best meet the objectives of their class.

This study adds to the data regarding DWCF and fluency. Hartshorn, et al. (2010) reported no significant difference between the treatment and control group in terms of fluency, and Hartshorn and Evans (2015) saw no significant difference in the number of words written from pretest to posttest using the original strategy. When Eckstein, et al. (2020) changed the DWCF process to an every-other-day writing schedule with timely feedback, they found the writing to have more words and clauses, and longer sentences. These results add new insight regarding how fluency can be increased with short, daily writing sessions by students with advanced proficiency.

Regarding complexity, Hartshorn, et al. (2010) observed a slight decline, Hartshorn and Evans (2015) observed no difference, and Eckstein et al., (2020) saw more complex writing in three lexical categories and three syntactic categories using an every-other-day schedule with timely feedback. The current study adds to the DWCF literature on complexity because no statistically significant difference was observed in either of the treatment groups using the MLTU or in the number of clauses per T-unit.

These mixed results could be explained by the trade-off hypothesis where accuracy and complexity compete for the learner's attention. Biber and Gray (2010) noted that clausal compression is one sign of increased complexity due to greater complexity at the phrasal level. Looking more closely at written complexity at the phrasal level may be a useful area for future study by analyzing student paragraphs such as those examined in this study through an online L2 syntactic complexity analyzer (see Ai & Lu, 2013; Lu, 2010, 2011; Lu & Ai, 2015).

Since both treatment groups used the DWCF process for the same amount of time each week, it would have been reasonable to expect that they would have demonstrated the same results. Nevertheless, the every-other-day group only saw improvement in accuracy, suggesting the tension between form (accuracy and complexity) as the trade-off hypothesis predicts (Skehan, 2009). On the other hand, the daily group experienced increases in both accuracy and fluency, suggesting student capacity to draw from multiple cognitive pools at the same time when driven by a more complex task (Robinson, 2003). It would seem that reducing the amount of time given to complete a writing task would make it less complex, but decreasing the time allotted

for the task may have been precisely what pressed students to produce their thoughts in writing within the given time.

We could also consider that reducing the task to 5 minutes may have relieved some emotional strain on students by limiting the amount of writing required, the volume of feedback received, and the number of corrections that needed to be fixed, thus reaching the optimum level of feedback. Cognitive resources may not have been so demanding that they were diverted to off-task or self-regulatory activities. As students gained more experience with the task type and with their knowledge of English, they seemed less affected by the negative emotions that may accompany high frequency feedback (Lam et al., 2011). The writing time of 5 minutes seemed to provide the most useable amount of feedback to support acquisition.

This DWCF strategy seems to be a fairly useful resource for grammar and writing practitioners. Teachers should be empowered to make research-based decisions for the use of DWCF in their classes based on their class schedules and objectives. Since the beliefs and practices of many L2 writing practitioners may not always be well aligned (Chen, 2023; Hartshorn, 2023), and many practitioners may be unfamiliar with the findings of relevant research, we encourage those involved in teacher preparation and in-service training to examine findings such as these with their teachers. From here, we look to build the body of research that identifies how teachers can use DWCF to assist their students in continuing to make gains in grammatical accuracy and also develop fluency.

Limitations and Future Research

A number of limitations should be considered. Because intact classes were used, there was no random selection of students. However, quasi experiments have been fairly effective in the past (Evans, Hartshorn, McCollum & Wolfersberger, 2010; Hartshorn et al., 2010; Hartshorn & Evans, 2012, 2015; Shadish et al., 2002). Despite efforts to ensure processes were as similar as possible, different teachers were used across groups which could have created a teacher effect.

Further research could involve other modifications, such as the effects of eliminating tally sheets and error logs, writing drafts at home, recycling or alternating prompts (as seen in Eddington, 2014). Research into how we measure complexity could provide a different lens with which we see our previous results and determine future measurement. The continued use of a single prompt for an entire week versus the use of a new prompt for each writing session could play a role in fluency development and is another area of future research. In any case, research that uses a larger sample of participants and collects a more robust amount of data would be welcome in continuing to build a solid bank of evidence that informs best practices of teachers.

Moreover, future research could also consider an area of growing interest—individual learner differences. This might include the effect of learning engagement in the

feedback process (see Mao & Lee, 2023; Shen & Chong, 2023; Zheng et al., 2023), which might help contextualize some of the observed inconsistencies across studies.

Conclusion

DWCF is a pedagogical strategy for teachers hoping to help their students improve the accuracy of their writing. It provides comprehensive feedback on writing that is produced within a short period of time, thus limiting the amount of feedback that needs to be provided by the teacher and processed by the learner. This study questioned the role of timing and feedback frequency within the DWCF process. Findings show that the feedback frequency for two treatment groups led to significantly greater accuracy. Students in the treatment group who wrote for 5 minutes every class period and received feedback daily also significantly increased their fluency along with gains in accuracy. This study adds to DWCF research in an important way as it is the first to find evidence of improvement in both accuracy and fluency at the same time in student writing based on a daily 5-minute writing limit. We encourage L2 writing practitioners to consider whether aspects of DWCF may be beneficial in their classroom contexts.

ORCID

https://orcid.org/0000-0002-0629-7410
https://orcid.org/0000-0002-5846-2385
http://orcid.org/0000-0002-3667-4571
https://orcid.org/0000-0002-4434-1558

Acknowledgements

Not applicable.

Funding

Not applicable.

Ethics Declarations

Competing Interests

No, there are no conflicting interests.

Rights and Permissions

Open Access

References

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). John Benjamins. https://doi.org/10.1075/scl.59.15ai

Akiyama, Y., & Fleshler, M. (2013). Effects of dynamic written corrective feedback on JFL students' homework. In K. Kondo Brown, Y. Saito Abbot, S. Satsutani, M. Tsutsui, & A. K. Wehmeyer (Eds.), *New perspectives on Japanese language learning, linguistics, and culture* (pp. 49–68). National Foreign Language Resource Center.

American Council on the Teaching of Foreign Languages. (2012). *ACTFL proficiency guidelines 2012*. https://www.actfl.org/sites/default/files/guidelines/ACTFLProficiency Guidelines2012.pdf

An, H. (2019). The efficacy of written corrective feedback on KFL learners' writing. *EALL Working Papers in Linguistics and Literature, 1*, 1-27.

Arroyo, D. C. & Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning, 68* (4), 942-972. https://doi.org/10.1111/lang.12300

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*(1), 2-20. https://www.sciencedirect.com/science/article/pii/S1475158510000020

Biber, D., Nekrasova, T., & Horn, B. (2011). *The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis* (ETS Research Report RR-11-05). ETS.

Bitchener, J. (2008) Evidence in support of written corrective feedback. *Journal of Second Language Writing 17*, 102–18. https://doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J. (2021). Written corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 207-225). Cambridge University Press.

Bitchener, J., & Knoch, U. (2008). The value of a focused approach to written corrective feedback. *Language Teaching Research, 12*, 409-431. http://doi.org/10.1177/1362168808089924

Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics, 31*(2), 193–214. http://doi.org/dng6fg

Bitchener, J., & Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207–217. http://doi.org/c62n2c

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

Boice, R. (1984). Contingency management in writing and the appearance of creative ideas: Implications for the treatment of writing blocks. *Behavior Research & Therapy*, 21, 537-543.

Casanave, C. P. (2007). *Controversies in second language writing: Dilemmas and decisions in research and instruction*. University of Michigan Press.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chen, W. (2023). Investigating novice EFL writing teachers' beliefs and practices concerning written corrective feedback across contexts: A case study from a complexity theory perspective. *Language Awareness, 32*(3), 465-486.

Cho, K. Y. (2021). A consideration of dynamic corrective feedback to develop writing accuracy. *언어과학연구* [*Language Science Research*] *97*, 273-287.

Collins, L., Halter, R. H., Lightbown, P. M., & Spada, N. (1999). Time and the distribution of time in L2 instruction. *TESOL Quarterly*, *33*, 655–668. https://doi.org/10.2307/3587881

Craven, L. (2023). *The effects of different types of unfocused corrective feedback on complexity, accuracy and fluency in L2 English academic writing* [Unpublished doctoral thesis]. University of Essex.

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Eckstein, G., & Bell, L. (2021). Dynamic written corrective feedback in first-year composition: Accuracy and lexical and syntactic complexity. *RELC Journal*. https://doi.org/10.1177/00336882211061624

Eckstein, G., Sims, M., & Rohm, L. (2020). Dynamic written corrective feedback among graduate students: The effects of feedback timing. *TESL Canada Journal, 37*(2), 78-102. https://doi.org/10.18806/tesl.v37i2.1339

Eddington, B. E. (2014). *A modified approach to the implementation of dynamic written corrective feedback*. [Unpublished master's thesis]. Brigham Young University.

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

Evans, N. W., Hartshorn, K. J., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System, 39*(2), 229-239.

Evans, N. W., Hartshorn, K. J., & Tuioti, E. A. (2010). Written corrective feedback: Practitioner perspectives. *International Journal of English Studies*, *10*, 47-77.

Evans, N. W., Hartshorn, K. J., Cox, T. L., & Martin de Jel, T. (2014). Measuring written linguistic accuracy with weighted clause ratios: A question of validity. *Journal of Second Language Writing, 24*, 33-50. https://dx.doi.org/10.1016/j.jslw.2014.02.005

Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in L2 writing pedagogy. *Language Teaching Research*, *14* (4), 445-463. https://doi.org/10.1177/1362168810375367

Ferris, D. (2003) Responding to writing. In B. Kroll (Ed.) *Exploring the dynamics of second language writing* (pp. 119–140). Cambridge University Press. https://doi.org/10.1017/CBO9781139524810.010

Ferris, D. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime ...?). *Journal of Second Language Writing*, *13*(1), 49–62. https://doi.org/10.1016/j.jslw.2004.04.005

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*(2), 181-201. https://doi.org/10.1017/S0272263109990490

Ferris, D., & Kurzer, K. (2019). Does error feedback help L2 writers? In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues (*pp. 106-124). Cambridge University Press. https://doi.org/10.1017/9781108635547.008

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, *10*(3), 161–184.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, *26*(2), 275–301. https://doi.org/10.1017/S0272263104262064

Gray, T. & Birch, J. (2000). Publish, don't perish: A program to help scholars flourish. *To Improve the Academy*, *19*(1), 268-284.

Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes, 38*, 1-13. https://doi.org/10.1016/j.jeap.2018.12.003

Hartshorn, K. J. (2023). Teacher cognition, pedagogical practice, and meeting student needs in L2 writing. In H. Mohebbi & Y. Wang (Eds.). *Insights into teaching and learning writing: A practical guide for early-career teachers* (pp. 71-86). Castledown.

Hartshorn, K. J., & Evans, N. W. (2012). The differential effects of comprehensive corrective feedback on L2 writing accuracy. *Journal of Linguistics and Language Teaching, 3*, 217-247. https://linguisticsandlanguageteaching.blogspot.com/2012/11/journal-of-linguistics-and-language.html

Hartshorn, K. J., & Evans, N. W. (2015). The effects of dynamic written corrective feedback: A 30-week study. *The Journal of Response to Writing, 2*, 6-34, https://journalrw.org/index.php/jrw/article/view/45

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*(10), 84-109. https://doi.org/10.5054/tq.2010.213781

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 461-473. https://doi.org/10.1093/applin/amp048

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp.2-21). John Benjamins Publishing Company.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. The National Council of Teachers of English. https://files.eric.ed.gov/fulltext/ED113735.pdf

Hyland, K. (2019). *Second language writing*. Cambridge University Press.

Hyland, K., & Hyland, F. (2019). *Feedback in second language writing: Contexts and issues* (2nd ed). Cambridge University Press.

Kamalian, A., & Lashkarian, A. (2014). The effect of dynamic written corrective feedback on Iranian elementary learners' writing. *International Journal of Applied Linguistics and English Literature*, 3(5), 47-56.

Karim, K., & Nassaji, H. (2019). The effects of written corrective feedback: A critical synthesis of past and present research. *Instructed Second Language Acquisition*, *3*(1), 28-52. https://doi.org/10.1558/isla.37949

Khalifa, N. M. E. A. (2022). Using online collaboration and dynamic written corrective feedback to enhance EFL secondary school students' writing skills. مجلة كلية التربية[*College of Education Journal*], *119*(4), 207-228.

Kulik, J.A., Kulik, C.-L.C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*, 79–97.

Kurzer, K. (2018). Dynamic written corrective feedback in developmental multilingual writing classes. *TESOL Quarterly*, *52*(1), 5-33. https://doi:10.1002/tesq.366

Lalande J. F. (1982). Reducing composition errors: an experiment. *Modern Language Journal*, *66*, 140–149.

Lam, C. F., DeRue, D. S., Karam, E. P., & Hollenbeck, J. R. (2011). The impact of feedback frequency on learning and task performance: Challenging the "more is better" assumption. *Organizational Behavior and Human Decision Processes, 116*, 217-228. https://doi.org/10.1016/j.obhdp.2011.05.002

Lavolette, E., Polio, C. & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. *Language Learning & Technology*, *19*(2), 50-68.

Lee, H. K., & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing, 24*, 307-330. https://dx.doi.org/10.1177/0265532297977299

Lee, I. (2003). L2 writing teachers' perspectives, practices and problems regarding error feedback. *Assessing Writing*, *8*(3), 216-237. https://doi.org/10.1016/j.asw.2003.08.002

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal, 63*, 13-22.

Lee, I. (2011). Working smarter, not working harder: Revisiting teacher feedback in the L2 writing classroom. *Canadian Modern Language Reviews, 67*(3), 377-399. https://doi.org/10.3138/cmlr.67.3.377

Lee, I. (2019). Teacher written corrective feedback: Less is more. *Language Teaching, 52*, 524-536. https://doi.org/10.1017/S0261444819000247

Lee, S. Y. (2009). *The efficacy of dynamic written corrective feedback on intermediate-high ESL learners' writing accuracy* [Unpublished master's thesis]. Brigham Young University.

Lim, S. C., & Renandya, W. A. (2020). Efficacy of written corrective feedback in writing instruction: A meta-analysis. *TESL-EJ, 24*(3), 1-26.

Liu, H., & Feng, M. (2023). The role of learner engagement with corrective feedback in EFL/ESL classrooms. *Frontiers in Psychology*, 14, 1-4. https://doi.org/10.3389/fpsyg.2023.1118467

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writer's language development. *TESOL Quarterly*, *45*(1), 36-62. https://www.jstor.org/stable/41307615

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, *29*, 16-27. https://doi.org/10.1016/j.jslw.2015.06.003

Mahmoud, E. M. (2023). Tertiary students' perceptions of dynamic written corrective feedback in an ESL classroom in the UAE. *Emirati Journal of Education and Literatures*, 1(1), 4-17.

Mao, Z., & Lee, I. (2020). Feedback scope in written corrective feedback: Analysis of empirical research in L2 contexts. *Assessing Writing*, *45*, 1-14. https://doi.org/10.1016/j.asw.2020.100469

Mao, Z., & Lee, I. (2023). Student engagement with written feedback: Critical issues and way forward. RELC Journal. https://doi.org/10.1177/00336882221150811

Marzban, A., & Arabahmadi, S. (2013). The effect of written corrective feedback on Iranian EFL students' writing. *Procedia-Social and Behavioral Sciences*, *83*, 1000-1005.

McQuillan, J. (2012, January 20). *All correction, all the time: Is written error correction worth the effort?* [Blog post]. http:// backseatlinguist.com/blog/?p=39

Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. Bloomsbury.

Nassaji, H. (2016). Anniversary article Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20(4), 535-562.

Netten, J., & Germain, C. (2004). Theoretical and research foundations of intensive French. *Canadian Modern Language Review, 60*, 275–294. https://doi.org/10.3138/cmlr.60.3.275

Oh, S., Bell, T. R., Hartshorn, K. J., Evans, N. W., & Damron, J. (2022). Efficacy of dynamic written corrective feedback on KFL writing. *Journal of Language Sciences, 29*(4), 311-341.

Ortega, L. (2012). Epilogue: Exploring L2 writing–SLA interfaces. *Journal of Second Language Writing*, *21*(4), 404-415. https://doi.org/10.1016/j.jslw.2012.09.002

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing, 21*(4), 375–389. http://doi.org/f4jt63

Quinn, P. G. (2021). Corrective feedback timing and second language grammatical development: Research, theory, and practice. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 322-340). Cambridge University Press.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105. https://core.ac.uk/download/pdf/77238704.pdf

Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science, 16*, 183–186. https://doi.org/10.1111/j.1467-8721.2007.00500.x

Sayad Deghatkar, V., Khodareza, M. R., & Valipour, V. (2022). The impact of dynamic written corrective feedback on the accuracy of English passive voice usage in foreign language narrative writing. *Biannual Journal of Education Experiences, 5*(1), 173-189.

Sayad Deghatkar, V., Khodareza, M. R., & Valipour, V. (2023). The impact of dynamic written corrective feedback on Iranian EFL learners 'writing accuracy, fluency and complexity. *Journal of Studies in Learning and Teaching English, 12*(1), 39-60.

Serrano, R., & Muñoz, C. (2007). Same hours, different time distribution: Any difference in EFL? *System*, *35*(3), 305–321. https://doi.org/10.1016/j.system.2007.02.001

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

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. https://doi.org/10.1016/j.system.2009.09.002

Shelley, A. (2014). *Dynamic written corrective feedback: Achieving manageability* [Unpublished master's thesis]. Brigham Young University.

Shen, R., & Chong, S. W. (2023). Learner engagement with written corrective feedback in ESL and EFL contexts: a qualitative research synthesis using a perception-based framework. *Assessment & Evaluation in Higher Education, 48*(3), 276-290.

Shintani, N. & Aubrey, S. (2016). The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *Modem Language Journal*, *100*(1), 296-319. https://doi.org/10.1111/modl.12317

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510-532. https://doi.org/10.1093/applin/amp047

Spada, N., & Lightbown, P. M. (1989). Intensive ESL programmes in Quebec primary schools. *TESL Canada Journal, 7*, 11–32. https://doi.org/10.18806/tesl.v7i1.557

Storch, N. (2021). Written corrective feedback and learners' objects, beliefs, and Emotions. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 581-597). Cambridge University Press.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327–369. http://doi.org/bd55rq

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. https://doi.org/10.1016/S1060-3743(99)80124-6

Truscott, J. (2001). Selecting errors for selective error correction. *Concentric: Studies in English Literature and Linguistics, 27*(2), 93–108.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255–272. http://doi.org/b8f8dd

Truscott, J. (2020). The efficacy of written corrective feedback: A critique of a meta-analysis. [Unpublished manuscript]. National Tsing Hua University. https://www.researchgate.net/publication/341453398_The_Efficacy_of_Written_Corrective_Feedback_A_Critique_of_a_Meta-analysis

Truscott, J. (2022). Efficacy of written corrective feedback in writing instruction: A critique of another meta-analysis. https://www.researchgate.net/publication/366424654

Truscott, J., & Hsu, A. Y-P. (2008). Error correction, revision, and learning. *Journals of Second Language Writing*, *17*(4), 292–305. http://doi.org/d959rq

Utsi, M. (2023). *Written corrective feedback for grammatical accuracy: The role of writing and feedback in language learning–a qualitative study of four teachers' beliefs*. [Unpublished master's thesis]. Umea University.

van Beuningen, C. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, *10*, 1-27. http://doi.org/10.6018/ijes.10.2.119171

van Beuningen, C. (2021). Focused versus unfocused corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 300-321). Cambridge University Press.

van Beuningen, C., De Jong, N. H. & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1-41. http://doi.org/10.1111/j.1467-9922.2011.00674.x

van Beuningen, C.G., DeJong, N.H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learner's written accuracy. *International Journal of Applied Linguistics*, *156*, 279-296. http://doi.org/10.2143/ITL.156.0.2034439

Wigglesworth, G. (2008). *Measuring accuracy in second language performance*. [Paper presentation]. TESOL International Convention, New York.

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

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

Zheng, Y., Yu, S., & Liu, Z. (2023). Understanding individual differences in lower-proficiency students' engagement with teacher written corrective feedback. *Teaching in Higher Education*, *28*(2), 301-321.