

Exploring Code-Based Feedback Options and Processing Patterns of L2 Writers: A Focus on Dynamic Written Corrective Feedback

Exploración de opciones de retroalimentación basadas en códigos y los patrones de procesamiento de los escritores de L2: Un enfoque en la retroalimentación correctiva escrita dinámica

Christina Torres
University of Central Florida (EE. UU.)

Florin M. Mihai
University of Central Florida (EE. UU.)



VOL. 6 (2025)

ISSN 2952-2013

<https://doi.org/10.33776/EUHU/linguodidactica.v6.8847>

Editorial
Universidad
de Huelva



Abstract:

Dynamic Written Corrective Feedback (DWCF) is a valuable approach for language instructors and students because it focuses on principles of learner needs as well as manageability, meaningfulness, timeliness, and constancy (Evans, et al., 2010). This study is a qualitatively analyzed protocol-analysis study which investigated patterns that intermediate and advanced English for Academic Purposes (EAP) students demonstrated as they processed and applied DWCF using editing or color feedback codes in individualized tutoring sessions. To accomplish the study's goal, eleven participants engaged in DWCF tutoring sessions conducted in a live online setting and recorded with shared screen capture. Tutoring sessions were evenly spaced at three-week intervals over a 16-week semester. Concurrent verbal report data (Ericsson & Simon, 1993) collected from forty sessions was transcribed and analyzed qualitatively using Storch and Wigglesworth's (2010) language-related episodes (LREs) as a guide. Verbal report data was triangulated with concurrent screen capture actions to determine the extent of participant engagement with the feedback provided as well as its application towards DWCF activities like charting feedback and editing paragraphs. This article presents findings for patterns regarding language form-focused LREs, such as verb tenses, article use, and prepositions. Findings revealed that participants applying DWCF with color codes generally had more patterns of extensive engagement compared to participants using editing codes feedback. However, extensive engagement was not necessarily paired with expected resolutions. Findings in this study support continued use of editing codes for DWCF as a meaningful and manageable option and presents some questions about feedback tracking charts within DWCF.

Keywords:

Language learning; Second Language Instruction; Writing (composition); Learning Processes; Feedback (learning)

Resumen:

La Retroalimentación correctiva escrita dinámica (DWCF) es un enfoque valioso para instructores de lenguas y sus estudiantes porque se enfoca en principios de las necesidades de estos, a la vez que manejabilidad, significatividad, puntualidad, y constancia (Evans et al., 2010). El estudio usó protocolo verbal analizado cualitativamente como metodología para investigar patrones que demostraron estudiantes en un nivel intermedio y avanzado en un curso de *Inglés para Propósitos Académicos* (EAP) mientras procesaban y aplicaban DWCF usando códigos editoriales o códigos de colores durante sesiones individuales de tutoría. Para cumplir esta meta, once participantes tomaron parte en sesiones de DWCF realizadas en vivo en línea que fueron grabadas con pantallas compartidas. Las sesiones fueron espaciadas uniformemente a intervalos de tres semanas en un semestre de 16 semanas. Los informes verbales concurrentes (Ericsson & Simon, 1993) fueron recogidos durante cuarenta sesiones y analizados cualitativamente usando episodios relacionados con lenguaje (LRE), teniendo en cuenta a Storch y Wigglesworth (2010) como guía. Los datos del informe verbal fueron triangulados con acciones concurrentes observadas en las grabaciones de captura de pantalla para determinar el alcance del compromiso con la retroalimentación proveída y su aplicación hacia las actividades de DWCF como completando una gráfica de la retroalimentación y corrigiendo los párrafos. Este artículo presenta patrones de LREs con enfoque en formas gramaticales como tiempos verbales, uso de artículos y preposiciones. Se encontró que participantes, aplicando DWCF con códigos de colores en general, tenían más patrones de compromiso continuo comparado con participantes usando códigos editoriales para retroalimentación. Sin embargo, el compromiso continuo no necesariamente cuadraba con resoluciones esperadas de LREs. Resultados de este estudio apoyan el uso continuo de códigos editoriales como una opción significativo y manejable en DWCF y presenta algunas preguntas sobre las gráficas de retroalimentación usados en DWCF.

Palabras claves:

Aprendizaje de idiomas; Enseñanza de una segunda lengua; Escritura (composición); Proceso de aprendizaje; Retroalimentación (aprendizaje)

Fecha de recepción: 29 de marzo de 2025

Fecha de aceptación: 27 de septiembre de 2025

Contacto: christina.Torres@ucf.edu

1. Introduction

Research on improving linguistic accuracy in second language (L2) writing has included a continued focus on the impact of written corrective feedback (WCF) (Bitchener, 2019; Bitchener & Knoch, 2010; Chandler, 2003; Ferris & Hedgecock, 2023). There are many types of WCF strategies, including direct, indirect, metalinguistic, and reformulation (Ellis, 2009). WCF for L2 writers is generally desired by students and helpful when thoughtfully applied, but there has not yet been agreement on what type of corrective feedback is most effective (Ellis, 2009; Ferris, 2011; Ferris & Kurzer, 2019). DWCF emerged as a response to the lack of consensus on best practices in WCF scholarship by refocusing feedback on two key principles (Evans et al., 2010). The first principle is a focus on learner needs, achieved by using short student-produced writing and individualized WCF. The second principle is that WCF must be “manageable, meaningful, timely, and constant” (Evans et al., 2010, p. 452).

Focusing on DWCF principles and learners’ needs provides practitioners and researchers an opportunity to address corrections systematically without overwhelming students or instructors. DWCF generally includes 10-minute writings, coded feedback, error tracking charts, error logs, and revision cycles, but it is adaptable in nature to meet classroom needs (Evans et al., 2010). One such adjustment made in DWCF research is the decision to adjust the number of revision cycles or rounds used with a class (Hartshorn et al., 2023; Hartshorn et al., 2010; Kurzer, 2018a). Studies of the impact of DWCF on student-produced writing have found that the practice can significantly increase the linguistic accuracy of this student-produced writing (e.g., Hartshorn et al., 2023; Hartshorn et al., 2010; Hartshorn & Evans, 2015; Kurzer, 2018a). Still, an emphasis on product-based research (learner output/writing) is one piece to the puzzle, and more research is needed to better understand how feedback uptake happens. The present study followed Bitchener and Storch’s (2016) recommendation on conducting cognitively informed research on WCF for L2 development by focusing on feedback processing to better understand how learners interpreted and used DWCF.

2. Theoretical Framework

Acknowledging the complexity of L2 learning, a premise of instructed L2 learning is that students improve accuracy with teacher guidance and practice. Skill Acquisition Theory (SAT) is a cognitive-based Second Language Acquisition theory that proposes automatization of rule-based knowledge into more easily accessible and possibly automatized knowledge - the movement from declarative to procedural knowledge (DeKeyser, 2007). Critiques of this theoretical perspective include the relatively vague concept of “practice” and an implication that L2 knowledge begins as declarative metalinguistic knowledge (Ellis, 2008). The present study acknowledges SAT in DWCF because further research on how learners interpret and use the components of DWCF is needed to advance understanding the strengths and shortcomings of this feedback system, especially since DWCF heavily draws inspiration from SAT (Hartshorn & Evans, 2015).

The study allowed exploration of how feedback is processed with the intent to clarify learners’ use of declarative and procedural knowledge during practice prompted through DWCF. It looked for the presence of stages Bitchener (2019) proposed for a single WCF episode: Attention to form and motivation, attention input, attending to the gap between the learners’ output and the received input, understanding the input, analyzing the input against long-term

memory, hypothesis formulation and testing, decision making, and consolidation or repetition of this process.

3. Literature Review

3.1 *Written Corrective Feedback and Codes*

Ferris (2022) shared that L2 writing feedback is “arguably the most important pedagogical subtopic within the (sub)discipline of L2 writing” (p. 344). L2 writing feedback is not limited to WCF for linguistic accuracy. However, research literature in support of WCF’s efficacy for linguistic accuracy conducted major investigations on types of feedback and how much feedback to provide in L2 writing. Some seminal studies indicated that learners preferred coded indirect feedback even though simple indirect underlining appeared just as effective as the more time-consuming indirect coded feedback method (Chandler, 2003; Ferris & Roberts, 2001). There has also been evidence that teachers may deviate from the established coded indirect feedback delivery plan by providing direct feedback, but students are still able to use the feedback they receive, including incorrectly marked errors (Ferris, 2006).

Color-coded feedback is an underexplored option in WCF referenced in the second edition of *Treatment of Error in Second Language Writing* (Ferris, 2011) as an alternative to written editing codes, but academic articles about teacher-provided color-coded WCF for linguistic accuracy in ERIC, LLBA, and Science Direct databases were limited. Investigating automated writing evaluation, Cotos (2011) explored color-coded feedback as a strategy for providing rhetorical feedback on L2 student writing, but not linguistic accuracy. In a study of Spanish language learners, Valentin-Rivera and Yang (2021) found that WCF caught learners’ attention more, as measured by eye tracking, when presented with bold text, underlining or with a different font color, and recommended exploring efficacy of visual saliency techniques.

With multiple types of feedback available and the often time-consuming nature of instructor-provided WCF, investigating the relevance of different WCF strategies on L2 writing, such as DWCF, continues to be valuable to teachers considering manageable ways to address their students’ needs.

3.2 *Dynamic Written Corrective Feedback (DWCF)*

Studies on DWCF have investigated the impact of the feedback system on various measures of linguistic accuracy, complexity, and fluency, as well as rhetorical competence. In these studies, the DWCF group has significantly outperformed control groups in improvements on linguistic accuracy (Hartshorn et al., 2010; Hartshorn & Evans, 2012; Hartshorn & Evans, 2015; Kurzer, 2018a). More recently, Hartshorn et al. (2023) compared groups engaging in DWCF within their grammar class with 10-minute writing tasks who met every other day and DWCF tasks with 5-minute writing tasks who met every day Monday-Thursday and a control group without DWCF. The study found that both DWCF groups outperformed the control group in accuracy, and the DWCF group which met daily had significantly higher scores in fluency compared to the other two groups. Complexity was not significantly affected by DWCF. These findings continue to support DWCF as a tool towards L2 writing accuracy. Studies of student perceptions while engaging in DWCF compared to grammar class instruction with textbooks and peer

feedback have also suggested a positive student response to DWCF (Kurzer, 2018b; Kurzer, 2019). There is evidence to support DWCF as an effective strategy to support L2 writers' development. However, to date, what happens while students process (interpret and apply) DWCF practices has yet to be explored in detail.

Prior DWCF studies have used variations of editing codes feedback, short abbreviations and symbols associated with separate metalanguage references (Hartshorn et al., 2010; Kurzer, 2018a). Alternative feedback code types merit further exploration, especially since a great deal of the research literature on WCF has devoted so much attention to the study of various error code types and their level of focus (see Ferris & Kurzer, 2019). Types of codes used to deliver feedback to students have been featured in the WCF literature but remain underexplored in DWCF. One example of code adaptation in DWCF was Kurzer's (2018a) change from the original 20 codes to 16 codes, now grouped by global, local, and mechanical errors as an effort to make the feedback more meaningful. Color codes feedback has not yet been used in the context of DWCF. Considering the potential impact of color on visual salience (e.g., Valentin-Rivera & Yang, 2021), color codes in DWCF can provide additional information about how students interpret coded feedback. Color-coded feedback was selected for investigation as another possibility towards meaningful and manageable feedback in DWCF, since this feedback does not require memorization of typical editing codes symbols (Brown, 2010; Shvidko, 2015).

This study extended the research on DWCF by focusing on how DWCF was processed (interpreted and applied) within a round of practice and by describing participants' patterns between two types of feedback code strategies: editing codes versus color codes in this context.

3.3 Research Question

Given the problem of identifying patterns in students' processing and uptake during DWCF, this study was guided by the following research question:

What patterns do English for Academic Purposes (EAP) students show as they process DWCF over the course of a semester using editing codes versus color codes?

4. Methodology

4.1 Study Design

A protocol analysis study was used with a qualitative approach to analyze L2 writers' concurrent verbal reports as they engaged in DWCF. Verbal reports allowed investigation of processes otherwise not accessible without verbal report (Bowles, 2010a; Ericson & Simon, 1993). While studies analyzing individual processing with verbal reports (think-alouds) in L2 writing and feedback have used quantitative methodology to assess the data (e.g., Kim & Bowles, 2019; Yang, et al. 2020), applying a qualitative analysis of verbal report allowed this study to provide a rich description of how L2 writers process this feedback. This choice allowed exploration to expand understanding of how DWCF works and the impact of an adapted color-coded version on DWCF.

Concurrent verbal reports from sessions were used as the primary source of data for analysis. This followed the recommendation that these concurrent reports are most accurate to reflect processing during a task because the content of concurrent verbal reports is pulled from readily accessible short-term memory (Brown & Rodgers, 2002; Ericsson & Simon, 1993). The study design considered this need to reduce time between the action and reporting as well as additional principles in Brown & Rodgers' (2002) guide to introspective research such as considering the additional cognitive load in verbal report, avoiding usual conventions of conversation, consideration of the actions concurrent to the verbal reports in interpretation, and understanding that automatic processes cannot be verbalized. Metalinguistic explanations were not explicitly prompted during the think-alouds. The researcher acknowledges the presence of metalinguistic explanations as a part of verbal coding schemes for language-related episodes (e.g., Qi & Lapkin, 2001; Storch & Wigglesworth, 2010) and depth of processing (Kim & Bowles, 2019). Therefore, metalinguistic explanations naturally offered by participants as part of the verbal report without researcher prompting were noted in the analysis.

4.2 Participants

This study included eleven international students enrolled in an EAP class at a large southeastern university in the United States. All participants were on F1 international student visas, in their first academic year, and ages 18-20 at the time of data collection. There were three females and eleven males in the group. Their first languages included: Arabic (6), Gujarati (2), Urdu (1), Korean (1), and Vietnamese (1). At this university, students with Test of English as a Foreign Language Internet-Based Test (TOEFL iBT) scores of 68-79, International English Language Testing System (IELTS) scores of 5.5-6.0, or Pearson Versant scores of 50-68 qualified to enter the university with support from EAP classes.

Placement into the EAP classes as well as an initial de-identified diagnostic writing assessment evaluated by two EAP instructional colleagues determined that the students recruited into this study were at an intermediate to advanced level of English language proficiency. Participants' home countries included: India, Oman, South Korea, Pakistan, and Vietnam. The majority were from Oman (6 students). While participants were obtained through a convenience sample, the researcher only recruited students from her colleagues' classes rather than her own to reduce the potential impact of student-instructor power roles.

All participants engaged in four DWCF sessions spaced evenly in the 16-week semester. Six participants received feedback using color codes, while five received feedback using editing codes. One participant in the editing codes group was unable to verbalize enough think-aloud data for inclusion in the data about processing patterns during DWCF. Therefore, data from six participants receiving color codes and four participants receiving editing codes was used for the study's qualitative analysis.

4.3 Instruments

Two strategies for coding student errors within the context of DWCF were investigated: DWCF editing codes and color codes. The specific types of feedback addressed were based on previous coding schemes for DWCF (Hartshorn, et al., 2010; Kurzer, 2018a). Related errors were also grouped into larger feedback categories. For example, errors with subject-verb agree-

ment, verb form, and verb tense are all grouped under “verb” errors. Additionally, what previous DWCF studies listed as “awkward” and “unclear meaning” were combined into a single “multiple issues” code. This system simplified the color-coded feedback to seven colors. The adapted feedback codes and corresponding explanations provided to participants for the current study still addressed noun, verb, lexical, and mechanical errors. Table 1 is an excerpt comparing the editing and color feedback codes and explanations used for nouns & determiners.

Table 1. Excerpt of Feedback Codes for Comparison

Editing Codes			
Feedback Category	Code	Specific Feedback	Examples Note: RED = incorrect
Nouns & determiners	S/PL	Singular/Plural	- Countable nouns: - Singular: Person, chair, student - Plural: People, chairs, students - Noncount nouns: honesty, furniture - Incorrect: Honesties, furnitures
	D	Determiners (articles)	- Incorrect: I bought chair. - Correct: I bought a chair. - Correct: The chair in my room is comfortable. - Correct: The chairs in my room are comfortable.
Color Codes			
Feedback Category	Specific Feedback		Examples Note: RED = incorrect
Nouns & determiners	Singular/plural		- Countable nouns: - Singular: Person, chair, student - Plural: People, chairs, students - Noncount nouns: honesty, furniture - Incorrect: Honesties, furnitures
	Determiners (articles)		- Incorrect: I bought chair. - Correct: I bought a chair. - Correct: The chair in my room is comfortable. - Correct: The chairs in my room are comfortable.

Source: own elaboration

4.2 Procedure

The sessions resulting in data collection were offered as supplementary tutoring outside of the regular EAP classroom activities. Students were enrolled into a Canvas course created for this project to facilitate writing sample collection. Participants were asked to keep their Zoom screenshare enabled during the sessions after removing any personally identifiable information from their desktops. Once permission was obtained at the beginning of each session, the session audio and screenshare were recorded for later data analysis.

Session procedures aligned with DWCF following the steps outlined in prior studies like Harts-horn et al. (2010), but condensed the steps into one focused session. The sessions also included a concurrent think-aloud component (Ericsson & Simon, 1993) to capture participants’ processing while working with the feedback. The sessions were the participants’ first exposure

to DWCF practices and, therefore, offered an opportunity to examine in detail how students respond to the process of learning and applying DWCF.

Each session began with informed consent procedures. After obtaining consent, the session continued with training (or training reminder in the second, third, and fourth rounds) on how to think aloud. Think-aloud training was demonstrated using a multiplication task as per Ericsson and Simon's (1993) recommendation and Swain and Lapkin's (1995) study to avoid priming participants and to emphasize verbalization of thought processes themselves rather than explanation of the thought processes in English. The procedure involved rounds of DWCF completed using participant-created, 10-minute paragraphs written from prompts based on the TOEFL Internet Based Test (iBT) independent writing prompts.

At each session, the participant would:

1. Write a 10-minute paragraph on a provided prompt using the Canvas quiz function
2. Receive immediate coded feedback from the researcher and a coded feedback reference chart
3. Think aloud while using a chart to track types and number of corrections in provided feedback
4. Think aloud while editing the paragraph using the feedback provided
5. Answer brief retrospective questions after the concurrent think-aloud tasks

All spoken data was audio recorded to allow for verbatim transcription. The participants' editing and charting process (steps 3-4 above), was recorded using Camtasia screen casting software by TechSmith to supplement concurrent observational notes with steps 1-5. All screencast video capture cropped out the participants' webcams for confidentiality. Per the DWCF process, participants submitted their feedback tracking charts and edited paragraph documents after completing an initial round of feedback application. Steps 2 and 4 were repeated as needed until the paragraph was error free or the participant was unable to identify further steps for corrections independently with the coded feedback provided. This was typically accomplished in two passes. After the first feedback pass, the researcher indicated remaining errors in the paragraph with simple underlining as done in previous DWCF studies (e.g., Evans et al., 2010).

The fifth step was to ask for an immediate verbal retrospective report from participants about the session. There were two retrospective questions at the end of each session:

1. Do you have any additional thoughts you'd like to share about the activities today?
2. Is there anything you would do differently?

The researcher would also ask the following third retrospective question in cases where the participant fell silent for a while (despite the researcher prompting) or did something unexpected.

3. I noticed (insert observation) during (insert the task(s) portion when it occurred). Can you tell me more about this?

The first DWCF session in week three of the semester ended with a request for participants to share basic demographic information including nationality, first language, and years of English study. University Institutional Review Board procedures for ethics and informed consent were followed. Data collection did not require participants to share sensitive personal information, and steps were taken to ensure confidentiality.

4.5 Data Analysis

Transcripts of 40 sessions including think-aloud data and concurrent actions taken by students (as observed through Zoom screenshare) were created for data analysis. Initial data coding in this qualitative study used the Language-Related Episodes (LRE) guide as described in Storch and Wigglesworth (2010). First, transcripts were segmented into LREs to identify portions with an “explicit focus on language” (p. 333). Portions of the transcripts with requests for clarification of the DWCF process, technology-related questions, or researcher requests were noted in transcripts but not included as LREs. These portions of data not falling neatly into the LREs such as reading aloud, task clarification requests, and feedback clarification requests were coded as “other.” In these “other” data cases, the information was recorded and checked for patterns but analyzed separately from the main LREs. Peer analysis between the first and second author was performed for rating and coding of a random 10% sample of transcript data to check for interrater reliability which was found to be 95% simple agreement.

LREs focus on three main areas: Form (F-LRE), lexical (L-LRE), and mechanics (M-LRE). With guidance from Storch & Wigglesworth (2010), F-LREs focused on verb tenses, word forms, article use, prepositions, and word order. L-LREs focused on word meanings or searching for a word or phrase. M-LREs included verbalizations about spelling and punctuation issues. This article reports findings from patterns of F-LRE processing - the largest LRE grouping - to report within the limited space of this article.

After identifying the type of LRE, the next step in coding decisions was to evaluate whether the LRE addressed the researcher’s feedback. The LRE was then marked as resolved as intended (correctly or incorrectly) or if it was left unresolved. Finally, following the LRE guide by Storch and Wigglesworth (2010), the level of engagement with an LRE could be marked as extensive (e.g., including metalanguage) or limited (e.g., just reading the feedback). Triangulation of data from the transcripts to the original screencast recordings and/or the submitted written products was used to confirm the LRE resolutions.

Following the qualitative nature of the study, the data analysis attended to patterns and shifts within engagement and resolutions as well as cognitive load during the study and across DWCF coding types while remaining open to other trends as they emerged from data. Details of how transcripts were prepared and coded as well as examples can be found in Appendix A.

After analyzing the data collected from all 40 sessions, LRE information across all participants and sessions was compiled and tallied. For example, all F-LREs for the editing codes group were organized into one document which separated LREs by EE and LE. The same was done for the F-LREs for the color codes group, again separated by EE and LE. Group tallies from the documents and tallies from each participant for each session were logged on a spreadsheet which was then used to examine overall patterns within the data.

The next step was to tally and visually represent patterns of counts and proportions from this large amount of data on a macro level first before returning to the transcription data for associated detailed descriptions which were used to illustrate these patterns. It is important to note that participants exclusively addressed the feedback which was provided to them regardless of the type of feedback codes used. This finding is not charted in the results section but is described in the discussion. Patterns that emerged from the F-LRE data exploration are described below.

5. Results

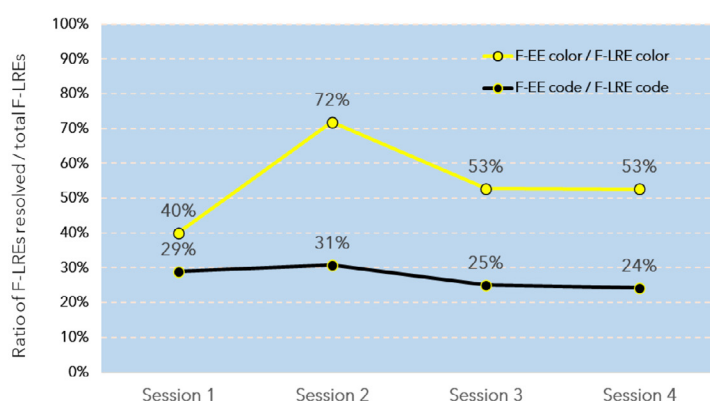
5.1 Limited Engagement (LE) and Extensive Engagement (EE) for F-LREs

The amount of engagement participants invested in F-LREs was relevant to statements about processing when using color versus editing codes. Raw LRE counts are not comparable across the colors and editing codes groups because of the uneven distribution of participants. The color codes group had 6 participants while the editing codes group had 4 participants with usable concurrent think-aloud data. Therefore, percentages were used to illustrate patterns across groups before returning to the transcript data to illustrate the patterns. EE compared to LE for all participants in all sessions was tallied and charted. For example, Participant 1 (P1) was in the color codes group and had two F-LREs in session 1. Of these, one was EE and one was LE. Therefore, P1 had 50% EE and 50% LE for the F-LREs in session 1. Participant 6 (P6) was in the editing codes group. P6 had seven F-LREs in session 1. One of these F-LREs was EE and six were LE. Therefore, P6 had 14% EEs and 86% LEs in the F-LREs for session 1. This procedure was followed for all participants across all sessions. The percentages allowed for viewing overall patterns across the editing codes and color codes groups despite the uneven participant distribution. Data shown in the charts that follow show these percentages (%). This visual representation using charts was done before returning to the transcript data to illustrate these findings further.

Figure 1 below illustrates patterns of the color codes group engaging in more EE compared to the editing codes group for F-LREs. "Color" is shorthand for color codes, while "code" is shorthand for editing codes feedback.

Figure 1

Proportion (in %) of F-EE to total F-LREs across four sessions



Source: own elaboration

When examining the transcript data on EE for participants in the color codes group more closely, this pattern of engagement matched the need to decipher feedback. Participants first needed to decide the intention behind color coded feedback before they could use it. This was observed especially in LREs that addressed actions for charting the feedback received. In contrast, the editing codes group as a whole did not use EE when making decisions about feedback charting. Examples of F-LREs with EE for charting decisions are illustrated in Appendix B.

Often, charting decisions in the color codes group aligned with the intention behind the feedback provided. One such example is illustrated in Appendix C. In some cases, participants in the color codes group created an incorrect mark in their feedback charts even though their verbalizations indicated an understanding of what the correction should be. Appendix D illustrates an example of a participant who understood the need to remove “that” from the example phrase in the feedback key, “Anyone that who wants,” but he logged the error as “insert something” even though the feedback referenced the need to “omit something.”

The color codes group was able to make decisions on editing by instinct by testing and checking editing decisions through reading and re-reading. However, occasionally, a participant went against a gut feeling on an editing correction in their paragraph because of an incorrectly charted piece of feedback earlier in the session. Appendix E shows an example of this type of decision.

The editing codes group almost exclusively used LE to chart the F-LREs, in contrast to the color codes group, which used mostly EE. Appendix F illustrates examples of this pattern.

The color codes group engaged in more deliberation for feedback chart activities which was illustrated in patterns of EE observed. Often, this deliberation in the charting activity led to more LE in later paragraph editing, possibly because participants had already spent time deciphering the feedback while charting. When the editing codes group used EE on F-LREs, it was for editing decisions in the paragraph which were paired with LE F-LREs for charting decisions. Examples of this pattern are available in Appendix G with an example from P7 reflecting on feedback in Session 1 from the sentence: “We faced (VT) many difficulties these days.”

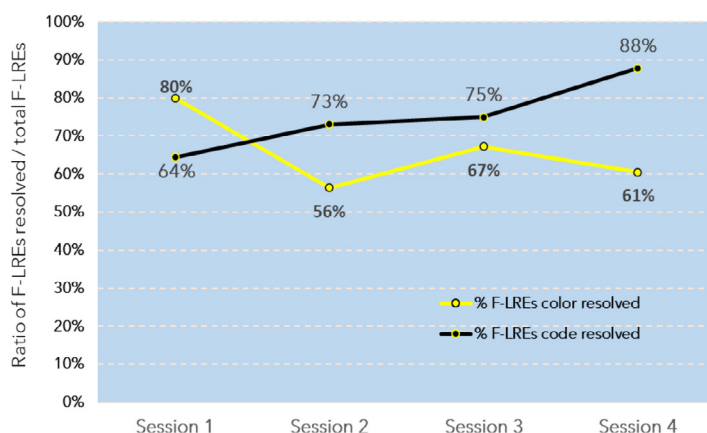
Participants across groups and sessions used reading and re-reading, guessing and checking correction options, metalanguage in problem-solving and checking the feedback key as strategies while interpreting and applying the feedback given during the DWCF sessions.

5.2 Resolutions and F-LREs

Resolutions to F-LREs that followed the intended resolution (i.e., expected grammar corrections to the paragraph or expected engagement with the feedback charting activity) were checked for each participant within each session. Visual representation of this data is found in Figure 2. General patterns revealed that the color code group initially resolved more F-LREs as compared to the editing codes group in session 1. However, the trend was for the editing codes group to resolve more F-LREs as the sessions progressed compared to the color codes group.

Figure 2

Proportion (in %) of F-LREs resolved to all F-LREs



Source: own elaboration

5.3 Proportions of Expected Resolutions Within EE and LEs for All LREs

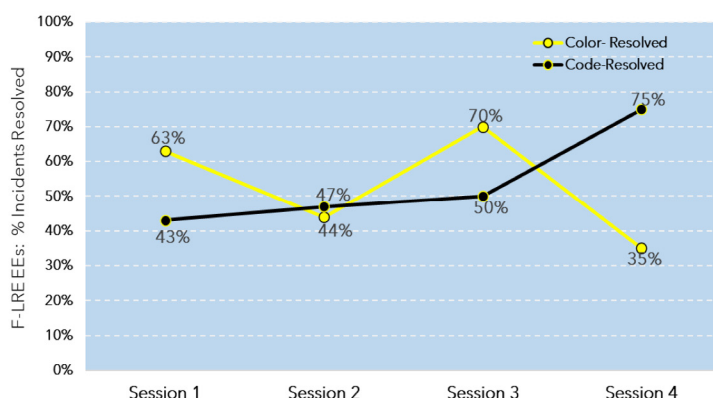
There were noticeable patterns within the overall data on EE and LE F-LREs as well as resolutions for F-LREs across color and editing code groups. Further exploration of patterns of resolutions within EE and LE engagements was needed for participants in the color and editing codes groups across all sessions because the goal of feedback provided is for it to be used as expected for the benefit of those using it. Information about resolutions, therefore, added a layer of information for discussion about what EE and LE meant for participants. Both counts and proportions of resolutions for EE and LE F-LREs are shared to provide context to overall trends used to interpret the transcript data. The interpretation of proportions should be made in connection to these tallies in Appendix H and the transcript data illustrating these patterns.

Figures 3 and 4 illustrate a general upward trend in EE and LE resolutions for the editing codes group. However, the EE and LE resolutions for the colors group did not show a clear pattern. The EE resolutions rose and fell in a zigzag pattern (Figure 3), while the LE resolutions fell and rose again to reach the same level of resolutions in Session 4 as in Session 1 (Figure 4). An example of a resolution that matched intended feedback can be found in Appendix I.

Results address three main points: First, the editing codes group tended to use the provided feedback towards more expected resolutions. Next, extensive engagement with an item of feedback was not necessarily always paired with expected resolutions. In fact, the color codes group using extensive engagement was also the group with less overall expected resolutions. This information about overall resolutions within engagement and the potential for misinterpretation of the color-coded feedback leads to the final points in the summary of results. First, the difficulty in interpreting color codes feedback paired with the charting activity supports the continued use of editing codes in DWCF. Additionally, further investigation is also recommended on the impact of feedback charts for efficacy in the DWCF system.

Figure 3

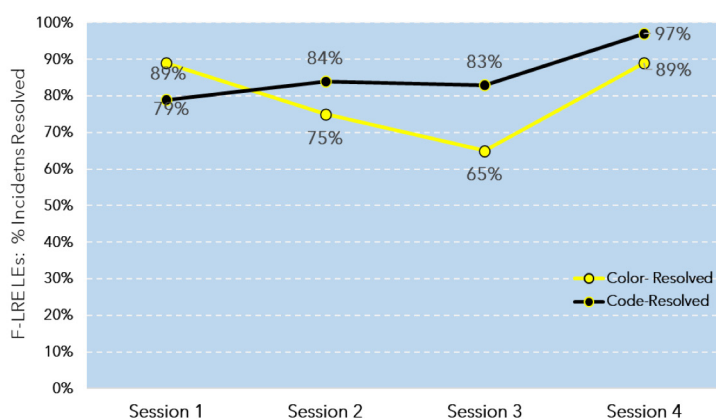
F-LRE EEs: % resolved



Source: own elaboration

Figure 4

F-LRE LEs: % resolved



Source: own elaboration

6. Discussion and Conclusions

This qualitative study investigated concurrent verbal reports to identify processing patterns within rounds of DWCF to better understand practice provided by the system. The study also investigated how an alternative feedback coding system, color codes, could work within DWCF. The research question was answered by analyzing concurrent think-aloud data and screen-capture from groups applying DWCF in individualized tutoring sessions across a semester. Generally, the editing codes group was able to maintain a more positive trend in their resolutions over the semester-long study compared to the color codes group. A more specific look at resolutions within extensive and limited engagement suggests that extensive engagement alone is not always indicative of an intended resolution with feedback.

The editing codes group consistently used a higher proportion of LE compared to EE in F-LREs throughout the DWCF sessions. The color codes group tended to use a higher proportion of EE. Nevertheless, EE with unexpected resolutions was noted in feedback charting tasks, and this ambiguity in charting later impacted editing choices and F-LRE trends. Also, using color codes required participants to recognize metalinguistic cues to apply these to charting and paragraph editing activities. Even some participants who could self-correct an error would sometimes mark a feedback chart incorrectly. In fact, confusion on how to use the feedback chart persisted throughout the entire semester. Some participants wrote words in the columns for feedback tallies in the later sessions of the semester even when the first columns in the chart were correctly marked with numbers. Some students marked the chart with numbers representing the order of feedback in the paragraphs: for example, 1 as the first error (regardless of the type), 2 as the second error, and so on. None of the participants discussed the feedback chart in retrospective questions at the end of their sessions. This finding of persistent charting difficulty across the semester leads to a recommendation to further examine the impact of charting as an element of DWCF. Future research to on the impact of the chart in DWCF aligns with Hartshorn et al.'s (2023) suggestion to examine the impact of separate aspects of DWCF.

In the study of patterns in DWCF feedback processing, patterns of F-LREs illustrated manageability and meaningfulness of the system. Color codes appeared to require a higher degree of metalanguage compared to editing codes to complete the DWCF charting activities, possibly because color codes appeared to be less intuitive. Editing codes appeared to be more meaningful and manageable for the participants in this study. Participants who needed guidance could identify the specific type of feedback associated with an editing code. Also, participants who already knew how to correct the errors noted by the editing codes feedback only needed to tally these symbols and insert a number in the feedback chart without referencing the feedback chart or key in great detail.

Practice is relevant for discussion here because Skill Acquisition Theory is one of the main theories referenced in discussions of DWCF (DeKeyser, 2007; Hartshorn & Evans, 2015). In the theory, practice is the way in which learners can progress from declarative to procedural or automatic knowledge. What happens during practice is relevant to further understand how DWCF works to improve language learners' linguistic accuracy.

DWCF principles emphasize manageable, meaningful, timely, and constant feedback (Evans et al. 2010). Participant engagement within LREs is important for this study of processing during DWCF. Students who understand and apply feedback successfully with limited engagement indicated an advanced understanding of the target issue, suggesting procedural knowledge, meaningfulness of the feedback system, or both.

The qualitative data obtained in this study supports that DWCF can encourage the presence of cognitive stages aligning with what Bitchener (2019) proposed for a single written corrective feedback episode. All participants in the study targeted the feedback provided regardless of the type of codes used in the DWCF sessions. The investigation of feedback code types and DWCF is valuable because it relates to meaningfulness, one of the major guiding principles in DWCF (Evans et al., 2010). Feedback can only be applied as intended towards linguistic accuracy when it is understood (e.g., Evans et al., 2010; Ferris & Kurzer, 2019). This study supports the continued use of editing codes with DWCF activities.

This study investigated how DWCF worked instead of asking whether DWCF effectively impacted linguistic accuracy in written products alone (e.g., Evans et al., 2011; Hartshorn et al., 2023; Hartshorn et al., 2010; Kurzer, 2018a) or asking about students' perceptions of DWCF (e.g., Kurzer, 2018b; Kurzer, 2019). Instead, the current study sought to better understand how participants processed DWCF in four DWCF rounds spaced evenly in one 16-week semester. This allowed the researcher to investigate what practice looked like for this DWCF system in terms of meaningfulness and manageability. Color codes was presented as an alternative feedback option for its potential impact on cognitive load and meaningfulness.

Findings in this study support the continued use of editing codes over color codes in DWCF to maintain meaningfulness for students. Participants already using procedural knowledge of language rules may have experienced confusion when trying to decipher declarative meta-linguistic rules to understand the nature of the given color-coded feedback. In contrast, the editing codes group was more able to effectively tally their feedback and apply it consistently with generally more consistent resolution patterns for F-LREs.

This article focused on F-LRE findings part of a larger research project that also included patterns for L-LREs, M-LREs, and other codes which emerged. Generally, F-LREs findings supported the continued use of editing codes in DWCF due to the patterns of limited engagement F-LREs paired with expected resolutions.

This study's data showed that some participants were able to make expected resolutions on editing decisions using the provided feedback but were unsuccessful in marking their feedback in the provided charts, which at times would misguide a later editing decision on a paragraph. Participants did not find the opportunity to compare their feedback using the chart intuitive. When used in classrooms, the charts must be explicitly explained, and their appropriate use must be confirmed before continued use. Misusing the charts for inaccurate tallies would invalidate their intended benefit of cross-writing comparisons of progress. Additionally, correcting student charts for accuracy may add an extra layer of work on teachers using the DWCF system which impacts the instructional manageability of the system. Educators should note that feedback can only be helpful when the students understand the feedback and what to do with it.

The practice in DWCF sessions highlights how application of feedback into multiple drafts is emphasized as student-produced writing guides the feedback provided. This emphasis on application aligns with previous research findings about the value of requiring revision to promote uptake on provided feedback (Chandler, 2003; Ekanayaka & Ellis, 2020). The finding that all feedback provided was attended to in paragraph editing throughout all sessions across this study supports the importance of revisions as an element in DWCF practice sessions. A take-away for educators is to ensure that students have opportunities to use their feedback towards future submissions in a class setting.

Interpretation of the findings should be made in the context of the small sample size within the exploratory qualitative study. The authors acknowledge that the sample size and uneven gender distribution do not allow for generalizability. Limitations in this study include taking DWCF from its natural classroom setting and placing it in an individual synchronous online tutoring setting to obtain the think-aloud data needed to examine processing. However, the methodology required this setting to obtain the data, which would otherwise not be possible in a classroom setting. Additionally, inviting participants to think aloud in their native languages would have been ideal, but translating for the five L1 backgrounds of participants was beyond the scope of this study. Future studies on DWCF processing may benefit from a larger sample

and the inclusion of a quantitative or mixed-methods approach to allow for more generalizable results.

References

- Bitchener, J., (2019) The intersection between SLA and feedback research. In Hyland, K. & Hyland, F. (Eds). *Feedback in second language writing: Contexts and issues* (2nd ed). (pp. 85-105). Cambridge University Press. <https://doi.org/10.1017/9781108635547>
- Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics*, 31(2), 193-214. <http://dx.doi.org/10.1093/applin/amp016>
- Bitchener, J. & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters. <https://doi.org/10.21832/9781783095056>
- Bowles, M. A. (2010a). *The think-aloud controversy in second language research*. Routledge.
- Brown, D. (2010, March). *Reshaping the value of grammatical feedback on L2 writing using colors*. Paper presented at the International TESOL Convention, Boston, MA.
- Brown, J. D., & Rodgers, T. S. (2002). *Doing second language research: An introduction to the theory and practice of second language research for graduate/master's students in TESOL and applied linguistics, and others*. Oxford University Press.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 students writing. *Journal of Second Language Writing*, 12(3), 267-296. [http://dx.doi.org/10.1016/S1060-3743\(03\)00038-9](http://dx.doi.org/10.1016/S1060-3743(03)00038-9)
- Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal*, 28(2), 420. <http://dx.doi.org/10.11139/cj.28.2.420-459>
- Dekeyser, R. (2007). Skill acquisition theory. In VanPatten, B. & Williams, J. (Ed.). *Theories in second language acquisition: An introduction*. (pp. 97-113). Erlbaum.
- Ekanayaka, W. I., & Ellis, R. (2020). Does asking learners to revise add to the effect of written corrective feedback on L2 acquisition? *System*, 94, 1-12. <http://dx.doi.org/10.1016/j.system.2020.102341>
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97-107. <http://dx.doi.org/10.1093/elt/ccn023>
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed). Oxford University Press.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. MIT Press.
- Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research*, 14(4), 445-463. <http://dx.doi.org/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(2), 229-239. <http://dx.doi.org/10.1016/j.system.2011.04.012>
- Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In Hyland, K. & Hyland, F. (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge University Press. <https://doi.org/10.1017/9781108635547>

- Ferris, D. (2011). *Treatment of error in second language student writing* (2nd ed.). University of Michigan Press. <https://doi.org/10.3998/mpub.2173290>
- Ferris, D. R. (2022). Feedback on L2 student writing: Current trends and future directions. In *Handbook of practical second language teaching and learning*. Routledge. <https://doi.org/10.4324/9781003106609>
- Ferris, D. R., & Hedgecock, J. S. (2023). *Teaching L2 composition: Purpose, process, and practice* (4th ed). Routledge. <https://doi.org/10.4324/9781003004943-1>
- Ferris, D., & Kurzer, K. (2019) Does error feedback help L2 writers? In Hyland, K. & Hyland, F. (Eds). *Feedback in second language writing: Contexts and issues* (2nd ed). (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. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- 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(1), 84-109. <http://dx.doi.org/10.5054/tq.2010.213781>
- 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(2), 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. *Journal of Response to Writing*, 1(2), 6-34. <https://scholarsarchive.byu.edu/journalrw/vol1/iss2/2>
- Hartshorn, K. J., Rice, S. H., Eckstein, G., & Evans, N. W. (2023). Dynamic written corrective feedback frequency and its effects on ESL writing fluency, accuracy, and complexity. *Feedback Research in Second Language*, 1, 7-32. <https://doi.org/10.32038/frsl.2023.01.02>
- Kim, H. R., & Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think-alouds. *TESOL Quarterly*, 53(4), 913-938. <https://doi.org/10.1002/tesq.522>
- Kurzer, K. (2018a). Dynamic written corrective feedback in developmental multilingual writing classes. *TESOL Quarterly*, 52(1), 5-33. <http://dx.doi.org/10.1002/tesq.522>
- Kurzer, K. (2018b). Student perceptions of dynamic written corrective feedback in developmental multilingual writing classes. *Journal of Response to Writing*, 4(2). 34-68. <https://scholarsarchive.byu.edu/journalrw/vol4/iss2/3>
- Kurzer, K. (2019). Dynamic written corrective feedback in a community college ESL writing class setting. In S. M. Anwaruddin (Ed.), *Knowledge mobilization in TESOL: Connecting research and practice*. Brill. <https://doi.org/10.1163/9789004392472>
- Qi, D., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10(4), 277-303. [http://dx.doi.org/10.1016/S1060-3743\(01\)00046-7](http://dx.doi.org/10.1016/S1060-3743(01)00046-7)
- Shvidko, E. (2015, May 1). *Written feedback: Using color-coded comments* [Web log post]. Retrieved from <http://blog.tesol.org/written-feedback-using-color-coded-comments/>
- Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition*, 32(2), 303-334. <http://dx.doi.org/10.1017/S0272263109990532>

- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391. <http://dx.doi.org/10.1093/applin/16.3.371>
- Valentin-Rivera, L., & Yang, L. (2021). The effects of digitally mediated multimodal indirect feedback on narrations in L2 Spanish writing: Eye tracking as a measure of noticing. *Languages*, 6(4), 159. <https://doi.org/10.3390/languages6040159>
- Yang, C., Zhang, L. J., & Parr, J. M. (2020). The reactivity of think-alouds in writing research: quantitative and qualitative evidence from writing in English as a foreign language. *Reading and Writing*, 33(2), 451-483. <http://dx.doi.org/10.1007/s11145-019-09970-7>

Appendix A

Coding Decisions

Table 2 shows a sample transcript excerpt with a coding decision from this study. This portion of data is from Participant 4's Session 4 and relates to the blue highlight indicating the need to include an article in the following sentence: "Due to current pandemic situation around the globe, online classes have become very common."

Table 2

Example of table structure of charted transcript during session 4

Speaker	Time	Spoken	Screenshare	Researcher Log	Coding
P4	3:47	So for the first one, I do not agree with this statement. Due to, the, current pandemic situation. Due to the, Oh! Due to the current pandemic. Right. Because it's specific and the one pandemic. [mumbles] Due to the current pandemic. And I guess that will be determiners, so this will be one.	Clicks back to the paragraph document. Types the creating Due to the current pandemic. Clicks back to the chart. Types a 1 in the determiners box for writing activity 4.	This is an accurate edit in the paragraph and marking in the chart. I've noticed that the folks using the color charts sometimes prefer to edit first and then mark the chart.	F-LRE TF Resolved correctly EE

Source: own elaboration

The transcript charts were divided by LRE, as illustrated in the example shown in Table 2, which represents one continuous LRE. In some cases, participants separated their charting actions in the sessions from their editing choices in the paragraphs. When this took place, these were treated as separate LREs. If a participant completed the identification of a feedback code along with charting and editing choices in one continuous stream of thought, this was treated as one LRE. Occasionally, the participants began an LRE on one area of feedback, paused to work on another area of feedback and different LRE before returning to the earlier LRE. In these cases, LREs were numbered so that the interrupted LRE could still be coded as one as recommended by Storch and Wigglesworth (2010). In the first DWCF session, participants more typically separated their charting activity from the paragraph editing activity because the researcher indicated the steps of the session in this way. However, as the semester continued, participants began

to combine the charting and editing activities, and the researcher chose not to intervene in the natural processes observed with the participants use of the DWCF system.

The example in Table 2 was a form-focused LRE (F-LRE) that targeted the provided feedback, was resolved correctly, and included extensive engagement. Storch and Wigglesworth’s (2010) LRE guide uses extensive engagement (EE) to describe an LRE with verbalized suggestions, counter-suggestions, explanations, or evidence of metalanguage. In this study, EE was also used to describe LREs where the participant made a direct reference to the feedback key which defined and provided examples of the feedback codes used in the DWCF rounds. This reference to the feedback key could be verbal or visual as captured by recorded screenshare. In contrast to EE, limited engagement (LE) was used to describe LREs which were shorter in length and included mere repetition of the feedback or an immediate charting and editing decision without deliberation.

The example in Table 2 reflected an F-LRE that was coded as resolved correctly because the participant’s editing and/or charting actions matched the intention behind the feedback or was a reasonable alternative (Storch & Wigglesworth, 2010). Unresolved and incorrectly resolved LREs were options available in the LRE coding guide. Additionally, a “resolved +-” code was developed during the qualitative data analysis for decisions that met one of two conditions:

1. The participant verbalized or completed an accurate editing correction but marked the feedback chart incorrectly.
2. The participant corrected the error targeted by the feedback but created a new error in the process of this editing.

Table 3 below illustrates an example of an F-LRE coded as “resolved +-” because there was an accurate spoken correction that did not match the editing changes in the document. This excerpt refers to the following written sentence including editing codes feedback from session 1 P11: “Telling someone small lies can protect (WC) the (D) fights...”

Table 3
Example of F-LRE coded as “resolved +-”

Speaker	Time	Spoken	Screenshare	Researcher Log
P11	9:36	Fights. D it’s ar- ticles. D. Telling small lies can avoid. Some fight. avoid a fight	Removes the before fights. Then, clicks back to the fee- dback chart and back to the paragraph. Types a before fights.	The first correction was accurate. Insertion of the article a created a new error. It’s interesting that she is saying the form correctly avoiding a fight but does not realize that the noun is written in its plural form.

Source: own elaboration

Appendix B

Extensive Engagement Examples from Color Codes Group

The examples in Table 4 below were taken from session 1 data, but these patterns of extensive engagement are consistent across all the sessions for the color codes group.

Table 4

Examples of F-LREs decision-making process

Participant	Time	Spoken	Screen Capture
P1	20:58	Should be blue for the chance. Mmm. It's a chance.	She mouses over the top blue portion of the chart key with explanations.
P1	21:12	I think I should put "a chance." like that. So maybe I should put an "a"	She scrolls back up to the feedback chart portion; then, scrolls back down to the explanations.
P1	21:25	determiners. Maybe one here.	She adds a 1 to the determiners box in the writing activity 1 column.
P4	7:08	Then I have to closed. That would be word form. That would be the correct form.	P4 clicks back to the word document and then to the feedback chart page.
P4	7:26	Beauty beautiful. So, I guess it would be here.	P4 mouses over the explanations for purple highlights. She adds the word closed to word form in the writing activity 1 column.
P5	2:28	[reading] being honest is a good way to understand each other's. Each other's. So, I think it's a singular, so we didn't have to put s. And others.	Types each other in red in the box within the feedback chart for singular/plural

Source: own elaboration

Appendix C

Color Codes F-LRE Charting - Matching Intended Feedback

Table 5 below references the F-LRE for the first blue mark as P4 in session 2 worked in reference to the sentence with feedback: "In **the** recent times, Internet has been our savior in so many situations."

Table 5

Example of F-LRE charting decision matching intended feedback

Session 2			
Participant	Time	Spoken	Screen Capture
P4	2:26	I do agree with the above statement. In the recent times. So, this blue, so it will be singular/plural or articles. The chair. I think the mistake is in the articles	Mouses over the sentence while reading. Pauses the mouse of the. Clicks back to the feedback chart. Mouses over the labels singular/plural and determiners (articles). Then, mouses over the feedback key in the examples for the determiners section.

Session 2			
Participant	Time	Spoken	Screen Capture
P4	2:51	In the recent times. The mistake is in the article, so should I have to change it or remove it? In recent times. An recent times. No. Are there any other articles? If I remove it. In recent times. In the recent times. In recent times. Okay, that sounds better. Ok, so I have to remove the.	Clicks back to the paragraph document. And back to the feedback chart. Clicks into the box for determiners in writing activity 2.

Appendix D

Color Codes F-LRE Charting - Not Matching Intended Feedback

The data excerpt in Table 6 below relates to the following brown feedback mark in P3 session 2’s paragraph: “Also, you can use it to contact **with** others easily.”

Table 6

Example of F-LRE charting decision not matching intended feedback

Session 2			
Participant	Time	Spoken	Screen Capture
P3	3:20	Incorrect. Anyone that who wants. Anyone who. I want see, I want to see. Insert. Maybe it will be insert.	Mouses over the examples for “omit” and “insert” something in the key while reading. Scrolls back up to the feedback chart. Types a 1 in the “insert something” box for writing activity 2

Source: own elaboration

Appendix E

Color Codes F-LRE Charting - Editing Confused by Charting

Table 7 below illustrates the participant making decisions about a sentence on the impact of grades on students’ motivation. The singular noun “student” should be plural to reflect a group of students, but P2 decides to leave the noun singular because of a reference to a prior charting decision to mark the blue feedback as an article rather than a singular/plural error stating, “No, it’s article.”

Table 7

F-LRE example of editing decision confused by charting decision

Session 3			
Participant	Time	Spoken	Screen Capture
P2	17:29	Now this one. Okay, let's go to the next one second. Here. Plural. No, it's article. In my opinion, it support, it's support, it supports the student. Because I'm basically just because I mentioned students here, I should write the. Okay. That it support the students. And, okay, because I mentioned that. I mentioned who's the students earlier here.	Mouses over student types a "the" before student.

Source: own elaboration

Appendix F

Editing Codes F-LRE Charting - Limited Engagement (LE)

The following examples in Table 8 were data excerpts from session 1. This pattern of using limited engagement for charting in the editing codes group was consistent across sessions.

Table 8

Examples of LE F-LREs from editing codes group

Participant	Time	Spoken	Screen Capture
P6	4:17	There's one S/PL and one article.	Types 1 into the singular/plural box and 1 for the determiners/articles box for writing activity 1
P7	8:01	Okay, the first one is we faced is a VT. So, one, two, three, four. I put it in the chart.	He mouses over all the VT symbols in the paragraph and then types the number 4 in the VT box for writing activity 1.
P11	4:54	Preposition, so I write one there.	P11 clicks back to the chart and types a 1 into the preposition box for writing activity 1

Source: own elaboration

Appendix G

Editing Codes F-LREs - Charting and Editing

Table 9 below shows an example of an LE F-LRE used in charting followed by an EE F-LRE used in editing decisions as participant 7 reflected on feedback in Session 1 from the sentence: "We faced (VT) many difficulties these days."

Table 9

LE F-LRE followed by EE F-LRE for feedback in editing codes group

Session 1				
Participant	Time	Spoken	Screen Capture	Coding
P7	8:01	Okay, the first one is we faced is a VT. So, one, two, three, four. I put it in the chart.	He mouses over all the VT symbols in the paragraph and then types the number 4 in the VT box for writing activity 1.	F-LRE LE
P7	10:11	Alright. Okay first of all is, we faced the we faced many difficulties. Is the VT is the main verb tense, right? okay. We faced. I think we face without -ed is in the present	P7 does not mouse over or do anything to the paragraph document.	F-LRE EE

Source: own elaboration

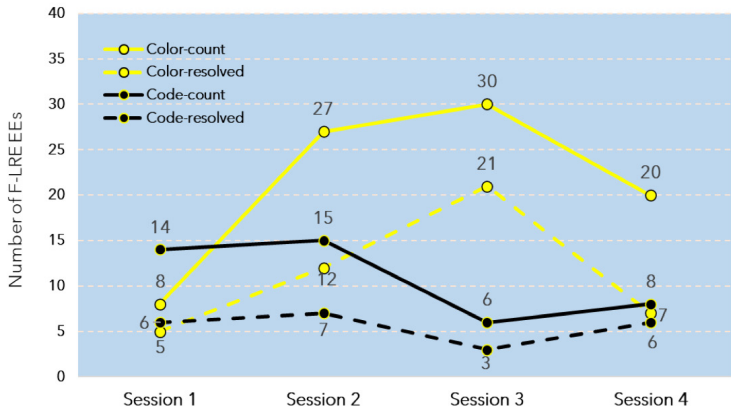
Appendix H

Counts of resolutions for EE and LE F-LRE across sessions

Figure 5 and Figure 6 below show the counts of F-LREs with intended resolutions (“resolved”) for the color (yellow line) and editing (black line) codes groups across all sessions. Figure 5 includes patterns of resolutions for F-LREs coded as EE, while Figure 6 illustrates patterns of resolutions for F-LREs coded as LE.

Figure 5

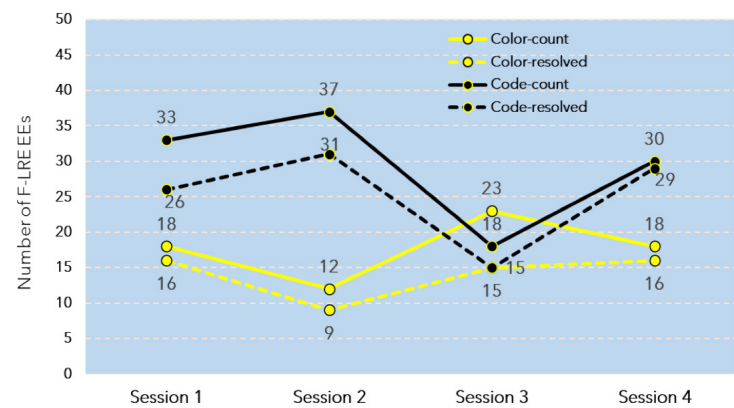
F-LRE EEs: Total vs. resolved counts for color codes and editing codes



Source: own elaboration

Figure 6

F-LRE LEs: Total vs. resolved counts for color codes and editing codes



Source: own elaboration

Figures 5 and 6 illustrate that there were more total counts than resolved counts for F-LREs for both color and editing codes groups across EE and LE, respectively. F-LREs marked as EE generally had a larger difference between the total counts (solid line) and resolved counts (dashed lines). These patterns suggest that EE did not always align with expected resolutions, especially for the color codes group. In contrast, Figure 6 showed a generally smaller difference between the total and resolved count lines for the F-LREs marked as LE. This pattern suggests that LE results for both the color codes and the editing codes groups engaged in more expected resolutions.

Appendix I

Example of F-LRE resolution matching intended feedback

This F-LRE is from Participant 7 in Session 1 and is about changing the verb tense from present simple “happen” to past simple “happened” in reference to past time.

Table 10

Example of F-LRE resolution matching intended feedback

Session 1				
Participant	Time	Spoken	Screen Capture	Researcher Notes
P7	36:25	You should say all the things things that happen. The things that happens. You should say you should say all the things that happens at the. No. You should say all the things that happened.	Clicks the cursor after the word happen. Adds an s to the end and deletes the s. Types a d at the end and edits to ed forming happened.	This is an accurate correction of the verb tense.

Source: own elaboration

Acknowledgements

Thank you to all the participants who made this study possible and to Dr. Aimee Schoonmaker for her help with participant recruitment.

Conflict of Interest Statement

The authors do not have a conflict of interest to disclose for this article submission.

Funding Statement

This study was completed without external funding.