Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 12 (4). 2022. 623-650
https://doi.org/10.14746/ssllt.2022.12.4.5
http://pressto.amu.edu.pl/index.php/ssllt

A classroom-based study on the effects of WCF on accuracy in pen-and-paper versus computer-mediated collaborative writing

Belén González-Cruz
University of Murcia, Spain
https://orcid.org/0000-0001-5345-7934
belen.gonzalez@um.es

Lourdes Cerezo
University of Murcia, Spain
https://orcid.org/0000-0001-5345-7934
lourdesc@um.es

Florentina Nicolás-Conesa
University of Murcia, Spain
https://orcid.org/0000-0001-8472-0811
florinc@um.es

Abstract

This study compared the effects of computer-mediated (CM) versus pen-and-paper (P&P) writing on written accuracy and feedback processing in tasks written and rewritten collaboratively following a pedagogical treatment in two intact authentic classrooms. The study involved 32 secondary education low-proficiency English-as-a-foreign-language (EFL) learners writing two descriptive texts collaboratively and receiving in-class training in the identification and correction of grammatical, lexical, and mechanical errors. Participants were provided with unfocused direct error correction (EC). Error logs were used to facilitate noticing of teacher corrections (i.e., feedback processing). Dyads were required to rewrite their texts for evidence of feedback uptake. Results indicate that writing collaboratively on the computer with the

availability of the Internet contributes to increased grammatical and lexical accuracy. No differences were found between writing environments regarding feedback processing or accuracy of rewritten texts.

Keywords: computer-mediated collaborative writing; feedback processing; L2 accuracy; pen-and-paper writing; written corrective feedback

1. Introduction

Research on written corrective feedback (WCF) addressing its effectiveness for second or foreign language (L2) development (e.g., Bitchener & Storch, 2016) has focused primarily on traditional, pen-and-paper (P&P) writing environments. However, educational technology is transforming the implementation of writing and feedback tasks in L2 classrooms, with learners making increased use of digital devices and Internet-based applications to write and revise their texts collaboratively.

Collaborative writing (CW), defined as the coauthoring of a single text by two or more writers, where the coauthors are involved in *all* stages of the composing process and have a shared ownership of the text produced (Storch, 2013), has been widely implemented in L2 classrooms (e.g., Storch, 2013) and substantially researched over the last decades (see Elabdali, 2021, for a review). This research has focused on: the collaborative dialogue or languaging (i.e., how learners verbalize linguistic problems; cf. Swain, 2006) that student writers engage in while jointly composing their text; the relationships formed in dyads/groups and how these affect the quantity and quality of the language-related episodes (LREs)[1] identified in collaborative dialogues; and the impact of LREs on the overall quality of joint texts. However, the value of teacher feedback as a trigger of attention to form and of noticing processes (e.g., Schmidt, 1990, 2001) has been insufficiently explored (see Storch, 2018; Zhang & Plonsky, 2020, for reviews).

The study reported in this paper aimed to fill this gap by inspecting effects of WCF in computer-mediated (CM) versus pen-and-paper (P&P) writing on written accuracy and feedback processing in tasks written and rewritten collaboratively following a pedagogical treatment. In what follows, we provide an account of relevant theoretical tenets and empirical findings regarding the pivotal elements of our study: pen-and-paper versus computer-mediated collaborative writing (CMCW) and WCF research.

[1] LREs are "any part of the dialogue in which the students talk about the language they are producing, question their language use, or other- or self-correct" (Swain, 1998, p. 70).

2. Literature review

2.1. Pen-and-paper versus computer-mediated CW

Current developments in educational technology as well as increasing requirements for collaboration in academic contexts across disciplinary areas and education levels are transforming CW tasks into CMCW tasks (Storch, 2022). This increase in the use of CM writing tasks in educational settings, especially at university, can be explained by the authenticity currently attached to Internet-enabled writing by the education community and specifically in L2 classrooms (e.g., Zhi & Huang, 2021). In L2 classrooms, writing on the computer fits the well-supported notion of process-oriented writing (as is often the case in the English as a foreign language [EFL] curriculum in Spain, where our study was conducted) as easier access to text enables writers to proceed towards the final draft (Lee, 2004).

Research comparing P&P and CM L2 writing is rather limited, and inconclusive results have been obtained regarding the differential effects of writing in one or the other environment on writing processes and products, mainly because of methodological differences across studies. Part of this research indicates that writing with a word processor fosters global-level revisions to a higher extent than writing on paper (Cheung, 2015), including lexical choices. Notably, this finding comes from studies conducted before the spread of Internet applications such as Google Docs, which suggests that it is writing in the word processor itself and not the resources offered by the Internet that promotes higher-level thinking and linguistic processes. It has also been claimed (Cheung, 2015) that although P&P writing is considered superior to digital writing for the retention of language (e.g., grammar rules, word meanings, etc.), writing in Google Docs and similar software allows students to make use of additional resources (e.g., spell-checkers, online dictionaries, or grammar websites) that may potentially improve the linguistic quality of their texts. Another way in which P&P has been considered superior to CM writing is that writers are more careful during linguistic formulation (Chan et al., 2017), which may be due to the difficulty in making changes in handwritten texts (Vasylets et al., this SI). In contrast, Internet-supported L2 writing may be less careful because learners know they can more easily make changes to what they write as they write it. This may mean that L2 writers may make more errors in CM than in P&P writing as they draft their texts. Yet, it is also possible that L2 writers make use of Internet resources to deal with their doubts about grammatical, lexical, or mechanical issues in drafts or revised texts.

Most research on the use of Web 2.0 applications such as wikis or Google Docs (see Zhang et al., 2021 for a review) has investigated university students' patterns of collaboration while writing (e.g., Saeed & Al Qunayeer, 2020), the

relationship between quality of collaborative single-draft texts and collaborative patterns (e.g., Abrams, 2019), or the effects of peer feedback on text revisions (e.g., Alharbi, 2019). This research indicates that co-producing an L2 text in CM settings is valuable for L2 development. However, results are still inconclusive (Zhang et al., 2021). There is therefore a need to explore whether CMCW has the potential to attract attention to form while writing by examining form-focused writing tasks and different learner profiles including the analysis of non-adult populations in non-university settings.

Despite potential learning benefits, many schools cannot yet afford sustained access to computer facilities for their students. There are also potential benefits related to P&P writing as handwriting has been associated with increased activation in the brain areas important for memory and learning (Vasylets et al., this volume), which could result in language learning gains when writing, processing WCF, and rewriting a text. In this sense, CM WCF also offers affordances for developing writers, such as the clear identification of written errors due to the greater legibility in the CM environment.

In view of the above, we may assume that L2 writers may employ their cognitive and linguistic resources differently depending on writing environment. Considering the scarce research that compares the effects of teacher WCF in P&P and CM writing (Pearson, 2022), especially with non-university students, it may be theoretically and pedagogically relevant to explore whether there are specific L2 benefits associated with writing and processing feedback on P&P or Internet-enabled devices. To fill this gap, the current study explored the effects of teacher WCF on the linguistic accuracy of texts written and rewritten collaboratively by adolescent low proficiency EFL learners after receiving and processing unfocused direct feedback in CM and P&P environments. In what follows we review relevant studies on WCF.

2.2. Research on teacher WCF

WCF is needed to help learners pay attention to and notice different features of language as attention to form and noticing are necessary for language learning, regardless of writing environment. Following Schmidt's (1990, 2001) noticing hypothesis, attention is what allows learners to attend to and notice the problems in their interlanguage (IL), namely, the differences between what they want to produce and what they can produce. Noticing must involve at least some level of awareness to result in successful intake (temporal retention of the correct use of language). However, a high level of awareness or understanding is needed to restructure the language system, which can lead to language learning (Schmidt, 1990). Following Leow's (2020) model of the L2 learning process

in instructed second language acquisition (ISLA), WCF, if attended to and noticed, allows learners to make the connection between their prior inaccurate knowledge or output and the correct L2 forms.

To date, there is no firm consensus as to the relative effectiveness of different types of WCF (see Karim & Nassaji, 2019 for a review). Research has also explored the factors that mediate its effectiveness, such as the cognitive and linguistic processes in which learners engage while processing (i.e., noticing, understanding, and applying) corrections (see Coyle et al., in press, for a review of feedback processing research). With some exceptions (Cánovas Guirao et al., 2015; Coyle & Roca de Larios, 2014, 2020; Wigglesworth & Storch, 2012b), most WCF research has been implemented in individual writing, but there is a need of research on CW and collaborative feedback processing, as these may result in more improvements in accuracy than individual writing. This may be so because processing teacher feedback collaboratively may encourage learners to engage more deeply with it (Storch, 2022) and facilitate noticing of L2 forms and of the gap between the learners' IL and the target language forms (i.e., the corrections) through collaborative scaffolding that may allow learners to understand WCF (Wigglesworth & Storch, 2012a).

Empirical research on the effects of WCF on CW is also limited. Wigglesworth and Storch (2012b) compared the effectiveness of reformulation and indirect EC with adult English-as-a-second-language (ESL) learners. They found that both led to improved accuracy in rewritten texts, with an advantage for the more explicit feedback (i.e., reformulation). In an EFL context, Coyle and Roca de Larios (2014) compared the use of models and direct EC with children. They found that students receiving direct EC incorporated more grammatical corrections and produced more acceptable and comprehensible rewritten texts overall, while those receiving models noticed lexical and grammatical features that were partially understood and partially incorporated in their texts, given the difference between the level of language offered in the model and their L2 proficiency. Cánovas Guirao et al. (2015) found that models were useful for promoting children's attention to lexis and chunks of language rather than to grammar. Coyle and Roca de Larios (2020) examined the use of models with children writing collaboratively in different instructional settings (EFL and CLIL – Content and Language Integrated Learning). Their results indicated that all children relied on the identification of surface differences between the model and their own draft. However, the CLIL pairs incorporated over half of the features noticed, while the EFL pairs incorporated only around 20%.

It follows from the above that more explicit types of feedback (e.g., direct EC, reformulation) appear to be more effectively noticed than less explicit types (e.g., indirect EC, models) and that, with younger and lower proficiency EFL

learners, even explicit corrections are not always noticed and understood. Research has also suggested that young learners could be trained to process feedback to notice the difference between their own writing and their teachers' corrections and thus maximize its effectiveness (e.g., Cánovas Guirao et al., 2015). One way of drawing learners' attention to linguistic features is by using consciousness-raising tasks. From a cognitive viewpoint, Ellis (2002) explains that these tasks engage learners in linguistic analysis, which may help them develop explicit knowledge. However, this may not be the case with low proficiency learners, who may otherwise benefit from collaborative dialogues. Collaborative dialogues can help learners understand feedback through co-constructed scaffolded support and guidance (DiCamilla & Antón, 1997) and can result in improved accuracy in subsequent rewritten texts (e.g., Watanabe & Swain, 2007). Within this sociocultural approach, Adair-Hauck et al. (2010) proposed the presentation-attention-co-construction-extension (PACE) instructional sequence for consciousness-raising activities based on guided induction and dialogic rule formulation.

The current study included a pedagogical intervention to maximize learners' noticing of error corrections, potential uptake, and accuracy in CW tasks using a three-stage spot-the-error activity (see Method) in which we followed the PACE sequence. By guiding learners in the reflection on their own errors we sought to foster higher depth of processing that may lead to awareness at the level of understanding (e.g., Leow, 2020), which is in turn supposed to foster IL development. This pedagogical intervention addressed Roca de Larios and Coyle's (2022) recommendation about giving clear guidelines to learners on how they are expected to use feedback to promote their engagement with metalinguistic (ML) awareness-raising activities and thus potentially contribute to language development in real classrooms.

2.3. A focus on authentic L2 classrooms in L2 writing and WCF research

Most WCF research has been implemented in controlled, laboratory-style conditions. The conclusions drawn from this strand of research have contributed to the maturity of the field, making it clear that one reason why there is no unequivocal conclusion about its effectiveness is that the instructional setting in which the research is implemented is also relevant. In other words, authentic L2 classrooms (with their authentic curricula and language learning, teaching, and assessment goals and criteria) may require a research methodology that differs from that employed in controlled studies. In this sense, Leow (2020) stresses the need to situate future WCF studies within the instructed setting.

This curricular approach to WCF research needs to address the integration of writing practice about different genres established in the curriculum with

other types of language practice in relation to error types and proficiency of the learners, among others (Leow 2020, p.110). Following these suggestions, our classroom-based study sought to inform pedagogical practice by including a metalinguistic awareness task about writing and feedback processing in addition to a pedagogical intervention about the characteristics of descriptive writing to help students improve their compositions in this genre. The study complied with the characteristics of ISLA applied WCF research proposed by Leow (2020) in that the data gathered from L2 writers were produced authentically (in real EFL classrooms), within the L2 syllabus, and over a period of time to simulate the language curriculum; in that both the teacher and L2 writers were involved in the process; and in that there was ecological validity to its findings.The following questions thus framed our research:

1. To what extent does the writing environment (CM vs P&P) foster increased accuracy in collaborative writing conditions?
2. To what extent does the writing environment affect the learners' collaborative processing of teacher WCF?
3. To what extent does the writing environment mediate collaborative uptake of WCF in terms of the learners' incorporation of error corrections in rewritten texts?

3. Method

3.1. Context and participants

Thirty-two EFL students from two secondary schools in Spain participated in this study (mean age: 13.5). Their level of English was A1 according to Oxford Placement Test (OPT) results (CM group: $M = 40.50$, $SD = 9.39$; P&P group: $M = 35.14$, $SD = 4.28$). Nine dyads completed the tasks on the computer (Google Docs) and seven on paper.

The two schools shared curricular guidelines for EFL writing instruction and their students had similar weekly exposure to English (four 55-minute sessions). The teachers, in conjunction with the research team, designed the instructional materials, implemented them similarly, and formed the dyads considering who the students wanted to collaborate with. They also ensured that the proficiency of dyad members was similar, to promote equal engagement in the tasks. The CM dyads were familiar with the use of Google Docs.

No initial significant differences were found between the CM and P&P dyads, either in L2 proficiency ($U = 82.50$, $Z = -1.66$, $p = .09$, $r = .42$) or in written accuracy

as measured in the percentage of global errors and types of errors (grammar, lexis, and mechanics) made in the pre-test. The pre-test consisted of a text that students wrote prior to intervention to evaluate joint writing ability as well as knowledge of how to write a descriptive text in terms of language, content, and organization. The types of errors of both groups of dyads were similar. The most frequent errors concerned grammar, mechanics, and lexis (see Table 1).

Table 1 Pre-test accuracy scores

Variables	CM dyads (*N* = 9)	P&P dyads (*N* = 7)	Mann Whitney *U* test			
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*	*Effect size
Duration (minutes)	28.11 (7.36)	24.71 (5.88)	26.50	-.53	.59	.13
Number of words	169.89 (36.30)	144.43 (34.28)	15.50	-1.69	.09	.42
% of errors	21.35 (6.79)	32.79 (14.16)	14	-1.85	.06	.46
% of grammatical errors	64.44 (9.23)	49.63 (14.88)	14	-1.85	.06	.46
% of lexical errors	13.24 (6.75)	14.17 (7.07)	30.50	-.11	.92	.03
% of mechanical errors	22.32 (12.77)	36.20 (16.02)	15	-1.75	.08	.44

Note. * Effect size = Z/√N

3.2. Data collection procedures

Data were collected in six 55-minute sessions over 21 days (see Figure 1). In Session 1, dyads wrote the pre-test, a 150-word descriptive paragraph about a place where they usually spend their summer holidays (see Appendix A). The CM dyads wrote in Google Docs but were not allowed to use other Internet resources (e.g., bilingual dictionaries) so that writing ability without external assistance could be assessed. However, Google Docs has a predictive typing engine that could have helped learners to write their texts. In Session 2 (Instruction 1), the teachers explained the type of language that characterizes descriptive writing (sensory details; feelings; similes and metaphors; third person for descriptions; use of conjunctions) and analyzed two descriptive model texts. In Session 3, dyads wrote the draft of a new 150-word text describing their bedroom (see Appendix A). The CM dyads were allowed to consult Internet resources such as Google Translate or dictionaries while writing their texts in Google Docs. In Session 4 (Instruction 2; see Appendix B), students were trained to process (i) feedback using a three-step metalinguistic awareness activity (i.e., spot-the-error) consisting in the teacher's modeling of error identification, correction, and provision of linguistic reasons; (ii) dyads' identification, correction, and explanation of errors; and (iii) individual identification, correction, and explanation of errors. The teacher selected the most frequent errors from students' drafts (e.g., subject-verb agreement, lexical choices, spelling, or punctuation) to train learners to better complete their error logs after receiving teacher feedback (see Session 5 below).

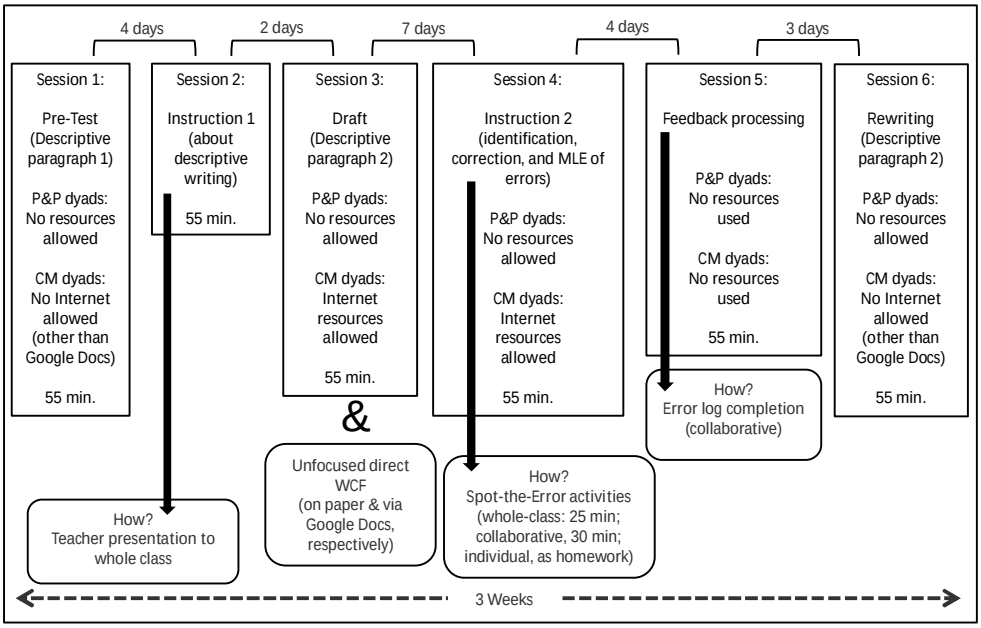

Figure 1 Data collection overview

In Session 5, dyads were given instructions to process teacher feedback collaboratively using an error log (see Appendix C). The teachers provided *unfocused* EC as this is the usual pedagogical practice in the Spanish education context. Additionally, the low proficiency of the participants was determinant in the choice of *direct* EC, which has been proven more effective for the improvement of accuracy in text revisions than indirect EC with low-level learners (Bitchener, 2019). The P&P dyads received the feedback on a photocopied version of their texts, whereas the CM dyads received asynchronous feedback using the track changes functionality of Google Docs. For the collaborative processing of errors through error logs, dyads were required to read their corrected drafts to notice the errors that had been corrected and explain the corrections collaboratively (see Appendix C). The CM participants were granted access to any Internet resources, such as online dictionaries. In Session 6, dyads were given their "clean" (i.e., uncorrected) drafts and were asked to rewrite them (see Appendix A) without having access to Internet resources other than Google Docs.

3.3. Data sources and coding

There were two main sources of data: written texts (i.e., pre-tests, drafts, and rewritten texts), and error log worksheets (i.e., collaborative feedback processing). To code errors, the more experienced researcher trained one of the authors, a doctoral

researcher. The former coded the CM data, and the latter coded the P&P data. Problematic instances in the two sets of data were discussed by the two researchers to ensure the coding schemes were applied consistently and as similarly as possible.

Errors in the written texts were coded using a scheme adapted from Nicolás-Conesa et al. (2019) (see Appendix D). A distinction was made between overall accuracy in terms of the number and percentage of errors made in drafts and rewritten texts, and uptake of WCF in relation to the percentage of incorporations of teacher corrections in rewritten texts. Uptake measures entail dyads' appropriation of the feedback received as evidenced by the changes made in rewritten texts (Storch & Wigglesworth, 2010). Uptake of WCF was coded following the scheme in Nicolás-Conesa et al. (2019) and Manchón et al. (2020). This analysis consisted in the comparison of drafts with rewritten texts to analyze whether errors had been successfully corrected, unsuccessfully corrected, uncorrected, or deleted. As for the error logs, we coded whether each error and its respective teacher correction had been successfully noticed, unsuccessfully noticed, or unnoticed by the two dyad members (see Appendix C).

3.4. Data analysis

A quantitative approach to data analysis was employed to facilitate comparison with previous WCF research. Overall accuracy was computed in terms of the error rate percentage for the three pieces of writing: (total number of errors/total number of words) × 100. The evolution of errors across tasks was also computed by subtracting the percentage of errors made in one text from the percentage of errors in the previous text. We also computed the percentage of grammatical, lexical, and mechanical errors (number of grammar, lexical, mechanical errors/total number of errors) × 100 in the pre-tests and in draft texts, as well as the number of specific error types within each of the three categories. For the rewritten texts, the overall percentage of uptake was computed: (number of errors (un)successfully corrected, uncorrected, or deleted/total number of errors) × 100, as well as the uptake of specific error types for grammar, lexis, and mechanics. Finally, regarding error logs, the overall noticing percentage was computed: (number of errors (un)successfully noticed, unnoticed/total number of errors) × 100, as well as the percentage of specific error types noticed.

Mann Whitney *U* tests were conducted to check for significant differences between the CM and P&P dyads in the noticing of errors through error logs, and in the accuracy of texts composed before and after pedagogical interventions (from pre-tests to draft texts and rewritten texts). Wilcoxon signed rank tests were run to compare possible differences in the percentage of errors over time within P&P and CM dyads.

4. Results

The first research question investigated the extent to which the writing environment (CM vs P&P) fosters increased accuracy in collaborative writing conditions. With this purpose, we compared the pre-test to the draft version of a new text within dyads and between dyads (CM vs P&P). The CM dyads significantly decreased their percentages of global errors (Wilcoxon signed rank test: Z = -2.66, p = .01, r = .63) and of grammatical and lexical errors (see Table 2), although the differences were not significant for each isolated type of error (grammatical errors: Z = -1.01, p = .31, r = .24; lexical errors: Z = -1.48, p = .14, r = .35). The grammatical errors reduced by the CM dyads were mainly rule-based errors (e.g., tense, verb form, or articles), while the lexical errors were related to the wrong choice of words. Finally, these dyads significantly increased their percentage of mechanical errors (Z = -2.19, p = .03, r = .52). Specifically, punctuation errors increased (34 versus 46) but spelling errors decreased (29 versus 13). In contrast, the P&P dyads did not significantly reduce their percentage of global errors (Z = -1.69, p = .09, r = .45), be it grammatical (Z = -.34, p = *.74*, r = .09), lexical (Z = -.17, p = .87, r = .05), or mechanical (Z = -0, p = .1, r = .0) errors.

Table 2 Evolution of global errors and types of errors from pre-test to draft

Variables	CM (N = 9)			P&P (N = 7)		
	Pre-test	Draft	Evolution	Pre-test	Draft	Evolution
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*))
% global errors	21.35 (6.79)	14.18 (3.69)	-7.17 (4.24)	32.79 (14.16)	25.70 (6.71)	-7.10 (10.15)
% grammatical errors	64.44 (9.23)	57.40 (18.78)	-7.03 (16.12)	49.63 (14.88)	48.25 (16.81)	-1.38 (13.46)
% lexical errors	13.24 (6.75)	9.41 (6.16)	-3.83 (6.95)	14.17 (7.07)	14.77 (6.64)	.61 (9.12)
% mechanical errors	22.32 (12.77)	33.18 (15.54)	10.86 (14.46)	36.20 (16.02)	36.98 (10.91)	.78 (14.98)

From a qualitative angle, we found that after Instruction 1 the descriptive texts produced by the CM dyads in their drafts included far more adjectives and that basic word order rules (e.g., adjective + noun) were more correctly implemented as well. There were also qualitative gains for the P&P dyads. Whereas their pre-test descriptions basically consisted of lists of the objects, furniture, etc., found in the place described, in the drafts they used more adjectives (albeit not always correctly), and one dyad even included a simile (e.g., "my bed . . . is like sleep in the clouds"), a characteristic of descriptive writing explained during Instruction 1. Also, the feelings that the place (their bedroom) evoked in the writers were described (adjectives such as "comfortable," "beautiful," "calm").

Comparing draft texts, in which the CM dyads could use Internet resources outside Google Docs (e.g., online bilingual dictionaries), differences were found in the percentage of global errors. The P&P dyads made significantly more global

errors than the CM dyads despite writing a similar number of words in the same amount of time (see Table 3). No significant differences were found between CM and P&P dyads in specific error types. The distribution of the most frequent types of errors was similar. The CM and P&P dyads made more grammatical errors followed by mechanical errors (spelling and punctuation) and lexical errors.

Table 3 Errors in draft texts

Variables	CM draft (*N* = 9)	P&P draft (*N* = 7)	Mann Whitney *U* test			
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*	Effect size
Duration (minutes)	33.78 (10.94)	27.43 (3.95)	18	-1.43	*.15*	.36
Length (number of words)	158.78 (14.48)	141.86 (25.34)	16	-1.64	*.10*	.41
% global errors	14.18 (3.69)	25.70 (6.71)	3	-3.02	*.00*	.76
% grammatical errors	57.40 (18.78)	48.25 (16.81)	24	-.79	*.43*	.19
% lexical errors	9.41 (6.16)	14.77 (6.64)	18	-1.43	*.15*	.36
% mechanical errors	33.18 (15.54)	36.98 (10.91)	31.50	0	*1*	0

The second research question examined whether feedback processing environment had any effect on the dyads' noticing of the corrections provided by their respective teachers. The CM dyads were granted access to Internet resources beyond Google Docs (see Method) to process feedback, but they reported in a questionnaire completed after feedback processing that they had not made use of them. No significant differences were found between processing feedback via Google Docs and on paper (see Table 4). The error logs revealed that the P&P and CM dyads successfully noticed most of the errors they had made in their drafts (i.e., approximately 80%), whereas unsuccessful noticing of errors was practically negligible, meaning that few errors passed unnoticed.

Table 4 Error noticing in error logs

Variables	CM (*N* = 9)	P&P (*N* = 7)	Mann Whitney *U* test			
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*	Effect size
% successfully noticed	80.30 (17.68)	80.78 (6.03)	29.50	-.21	*.84*	.05
% unsuccessfully noticed	2.87 (4.23)	1.23 (2.22)	25	-.79	*.54*	.19
% unnoticed	16.83 (14.99)	8.64 (6.39)	19.50	-1.27	*.21*	.32

Grammatical errors and corrections were successfully noticed to a much higher extent than lexical and mechanical ones. On the other hand, the unnoticed errors were mechanical errors, particularly in the case of the CM participants (see Table 5).

Table 5 Errors noticed per category in error logs

Variables		CM (*N* = 9)	P&P (*N* = 7)
		M (*SD*)	*M* (*SD*)
% Successfully noticed	Grammar	67.16 (22.26)	53.56 (20.10)
	Lexis	12.30 (7.77)	14.90 (7.25)
	Mechanics	20.54 (19.76)	31.54 (13.59)
% Unsuccessfully noticed	Grammar	44.44 (52.70)	14.29 (37.80)
	Lexis	0 (0)	0 (0)
	Mechanics	0 (0)	14.29 (37.80)
% Unnoticed	Grammar	10.19 (24.92)	23.81 (38.32)
	Lexis	0 (0)	8.33 (10.76)
	Mechanics	78.70 (38.44)	53.57 (42.99)

The last research question inspected possible differences between both environments in the incorporation of teacher corrections into rewritten texts. In terms of global errors, both the P&P and CM dyads successfully incorporated more than half of the corrections, with no significant differences. However, the P&P dyads incorporated significantly more unsuccessful corrections than the CM ones. About a quarter of the errors were similarly left uncorrected in both groups and only very few errors were deleted in both groups (see Table 6).

Table 6 Incorporation of corrections into rewritten texts

Variables	CM (*N* = 9)	P&P (*N* = 7)	Mann Whitney *U* test			
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*	Effect size
% Successful corrections	62.29 (24.83)	59.43 (15.62)	31.50	0	*1*	0
% Unsuccessful corrections	2.62 (3.55)	11.65 (7.10)	7	-2.64	*.01*	.66
% Uncorrected errors	32.96 (24.23)	25.31 (8.91)	22	-1.01	*.31*	.25
% Deleted errors	2.13 (5.02)	3.61 (4.13)	21.50	-1.22	*.22*	.31

The CM and the P&P dyads significantly reduced their errors from draft to rewritten texts (Wilcoxon signed rank test: CM – *Z* = -3.06, *p* = .00, *r* = .72; P&P – Z = -2.37, *p* = .02, *r* = .63). No significant differences were found in the percentage of errors in rewritten texts (see Table 7).

Table 7 Accuracy development across writing tasks

Variables	CM (*N* = 9)	P&P (*N* = 7)	Mann Whitney *U* test			
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*	Effect size
% Errors in pre-tests	21.35 (6.79)	32.79 (14.16)	14	-1.85	*.06*	.46
% Errors in drafts	14.18 (3.69)	25.70 (6.71)	3	-3.02	*.00*	.76
% Errors in rewritten texts	5.11 (3.64)	9.03 (3.78)	14	-1.85	*.06*	.46

Regarding the types of errors that learners corrected in the rewritten texts, the P&P and CM dyads incorporated (successfully or unsuccessfully) into their rewritten texts grammatical and, to a lesser extent, mechanical corrections. As for mechanical errors, the P&P dyads incorporated significantly more unsuccessful corrections (U = 18, Z = -2.10, p = .04, r = .53) than the CM ones. The P&P dyads also tended to incorporate more successful corrections in relation to mechanics (U = 14, Z = -1.85, p = .06, r = .46) (see Table 8).

Table 8 Incorporation of corrections per error type in rewritten texts

Variables		CM (N = 9)	P&P (N = 7)
		M (*SD*)	*M* (*SD*)
% Successfully corrected	Grammar	61.68 (22.83)	42.79 (26.77)
	Lexis	12.60 (11.44)	15.39 (7.98)
	Mechanics	25.71 (20.74)	41.82 (21.47)
% Unsuccessfully corrected	Grammar	33.33 (50)	56.87 (31.79)
	Lexis	11.11 (33.33)	28.57 (40.50)
	Mechanics	0 (0)	14.56 (18.46)
% Uncorrected	Grammar	33.88 (31.62)	50.80 (15.04)
	Lexis	3.99 (6.58)	11.95 (11.21)
	Mechanics	28.81 (29.49)	37.24 (17.43)
% Deleted	Grammar	3.70 (11.11)	35.71 (47.56)
	Lexis	0 (0)	0 (0)
	Mechanics	18.52 (37.68)	21.43 (39.34)

5. Discussion

Our results showed that CMCW (i.e., writing in Google Docs with the availability of Internet resources) led to increased grammatical and lexical accuracy as compared to writing on paper without external resources. However, no difference was found between writing environments (CM and P&P) in terms of feedback processing or accuracy of the rewritten texts. In what follows we discuss our main findings in terms of the evolution of L2 written accuracy as a function of the use of Internet resources; the effects of environment on feedback processing; and the incorporation of teacher corrections (i.e., feedback uptake) into rewritten texts.

5.1. Evolution of L2 written accuracy and the use of Internet resources

Our results indicate that using Internet resources during joint text formulation fostered accuracy and, more specifically, grammatical, and lexical accuracy, better than writing on paper. These findings are in line with those of previous research about the advantages of CM to improve the quality of written texts (Cheung,

2015; Li & Cumming, 2001). On the other hand, our results stand in contrast with other studies that indicate that CM writing may have negative effects on students' writing quality because of the difficulty in typing words in a computer (Pennington, 1996) or because learners may focus on surface-level features such as spelling (Owston et al., 1992). It should be remembered that our participants were low L2 proficiency learners, but they were used to writing in Google Docs. However, having fewer grammatical or lexical errors while writing using the Internet does not necessarily mean that learners who write in a CM environment have actually developed their L2. Further research should investigate if this additional advantage of the CM environment is later reflected in better accuracy in L2 production tasks.

Accuracy improvement in the CM dyads can be explained by the additional support provided by Internet resources together with other task-related factors (although these may apply to both P&P and CM writing) and the characteristics of EFL instruction. Regarding task-related features, content similarity between the pre-test and the draft/rewritten text (see Appendix A) together with procedural task repetition (e.g., Manchón, 2014) may have driven all learners' attention to lexical and grammatical features. In addition, EFL instruction in Spain focuses on grammatical and lexical accuracy over mechanical features, which may have also helped to direct learners' attention to grammar and lexis. Along these lines, our training in descriptive writing also emphasized the importance of using adjectives (lexis) and the third person singular (grammar) in descriptions. Accordingly, the CM dyads could have produced more grammatically and lexically accurate texts than the P&P dyads because they could use Internet resources to look up words in dictionaries and consult grammar rules on the Internet. To be more precise about the specific affordances for written accuracy of different Internet resources, it would have been useful to track the Internet searches made by these dyads while writing their texts. Owing to data protection policies, the research team was not granted access to this potentially sensitive information.

On the other hand, we also found that the availability of Internet resources was not beneficial for mechanics and, more specifically, for punctuation issues. The CM dyads increased their punctuation errors from the pre-test (f = 34) to the drafts (f = 46). As a result, the frequency of punctuation errors was similar in the P&P and CM drafts despite the latter using Internet resources. The CM dyads may not have noticed punctuation errors when writing their texts because Google Docs highlights spelling errors but not punctuation issues. Therefore, CM dyads may not have consulted Internet resources to solve potential punctuation problems in their drafts. In a way, it could be suggested that when the CM dyads resembled the P&P dyads the most (in the pre-test), they used more careful text formulation (Chan et al., 2017) and committed fewer mechanical errors than when they did have access to the Internet. It may be concluded

that the availability (or not) of resources in different writing environments may drive learners' attention to distinct L2 aspects while composing their texts.

Finally, the large percentage of mechanical errors of both groups could also reveal a lack of L2 writing practice. The Spanish secondary education EFL curriculum emphasizes the development of oral over written output and these learners' prior writing experience was mostly word- or sentence-based. These learners were not used to writing paragraph-long texts, which require higher use of punctuation marks.

5.2. Effects of environment on feedback processing

Our second research question asked about the extent to which the writing environment affects collaborative processing of teacher WCF. The CM dyads were given access to Internet resources for the feedback processing stage, but they reported not using them. Accordingly, their supposed advantages at this stage were neutralized and an effect of the writing environment on feedback processing was not found. This finding seems to indicate the importance of L2 learners' agency in writing to make use of available resources (Cerezo et al., 2019). Likely reasons for not making use of Internet resources might be, among others, the explicitness of the feedback (direct EC), insufficient time, or lack of strategic knowledge to use the available resources for feedback processing. From a methodological viewpoint, some questions remain open. For instance, it is unknown whether the CM dyads would have used Internet resources to process WCF under different conditions (individual writing, indirect feedback, focused corrections, higher proficiency, or longer time to discuss errors). Future research could use concurrent data (e.g., Leow, 2020) such as audio or video recordings to capture the interaction taking place in collaborative dialogues while writing and processing WCF and thus better understand the motivations behind some decisions taken by the learners as well as the actual products resulting from those decisions.

Our training program was based on feedback processing and on the characteristics of descriptive writing, but we did not train students to use the Internet specifically for text formulation and feedback processing. The CM dyads were accustomed to using Google Docs for their day-to-day schoolwork and the teacher confirmed that her students resorted to Google Translate or Word Reference in their English lessons. However, this knowledge may not have been sufficient to use Internet resources strategically. As indicated by Li and Cumming (2001), availability of computers on their own cannot result in positive changes in writing without adequate training.

Regardless of the writing environment, all participants successfully noticed around 80% of the errors in their drafts and the overall percentage of unnoticed

errors was reduced. These findings can be related to the explicitness of the feedback and of the instrument to process feedback collaboratively (i.e., error log) as well as to the effectiveness of the metalinguistic training. These were common factors to CM and P&P dyads. Further research could be conducted to isolate the effects of each variable. Processing feedback on the screen of a computer or on paper did not make any difference in the amount of noticing for our low L2 proficiency participants. Further research could investigate learners' depth of processing and if processing feedback on a screen using longer texts and during longer periods of time could be an additional burden.

To our knowledge, there are no previous studies on feedback processing in CM writing to compare our findings in this specific environment. The percentage of noticing in our study is higher than that reported by Cerezo et al. (2019) and Manchón et al. (2020) in relation to upper-intermediate EFL learners with a background in Linguistics. In these two P&P studies, the participants also received unfocused feedback and processed it using a languaging table in which they wrote errors, corrections, and explanations to those corrections. However, the participants in those studies did not receive training to process feedback and they wrote either a similar number of words to that indicated in the present study, individually and collaboratively (Manchón et al., 2020), or longer texts individually (around 300 words in Cerezo et al., 2019). It remains an empirical question if the percentage of successful noticing of errors would have been so high if our participants had received unfocused direct feedback in longer texts. Our findings may indicate the effectiveness of the metalinguistic awareness-raising activity (Instruction 2) in curriculum-based tasks, which are typically short.

Regarding the types of errors successfully noticed, grammatical errors were noticed the most frequently, followed by mechanical (basically spelling errors) and lexical errors in both groups. This distribution of noticing coincides with the most frequent types of errors made in students' drafts. On the other hand, the use of direct feedback could have contributed to focusing learners' attention on grammar, as suggested by previous scholars (Suzuki, 2012) as well as the higher presence of grammatical errors over mechanical ones in the ML awareness-raising activity.

5.3. Incorporation of corrections into rewritten texts

The third research question explored the extent to which the writing environment mediates collaborative uptake of WCF in the rewritten texts. Results showed that there was no effect of the writing environment because the potential advantages of the CM dyads were neutralized when they decided not to use the Internet to process feedback. The participants in both environments successfully

incorporated more than half of their respective teachers' corrections into their rewritten texts. Some studies with individual upper-intermediate EFL writers have reported lower percentages of successful incorporation of corrections after feedback processing in P&P without a training program to notice errors in a narrative task (e.g., Cerezo et al., 2019). The participants in Manchón et al. (2020) incorporated a slightly higher number of corrections, but the task was a decision-making writing activity, which could have made it more difficult to write the text and could have made the corrections more memorable. Our results indicate the effectiveness of training low proficiency EFL writers to reflect on errors using unfocused feedback, which could lead to increased accuracy in rewritten texts. We provided unfocused feedback in short texts, which means that the range of error types could be more limited than in longer texts and therefore the processing load may be manageable for errors to be noticed and later incorporated into their rewritten texts by our low proficiency participants, regardless of the writing environment. These findings are in line with Hartshorn and Evans (2015), who propose that unfocused feedback is effective and ecologically valid, but the number of corrections provided should be manageable for the learners. This can be done if feedback is provided in short texts (around 150 words), as the ones required in this study.

The percentage of successful noticing of errors was higher than the percentage of successful correction in rewritten texts in both environments. Around 20% of the errors successfully noticed were not incorporated into the rewritten texts, which indicates the lack of total correspondence between the number of errors noticed and corrections introduced into rewritten texts after receiving and processing unfocused feedback (Cerezo et al., 2019). Further research could investigate if more cycles of writing and training sessions to process feedback could lead to higher levels of noticing and higher percentages of accuracy. This research should be conducted using curriculum-oriented writing tasks in which learners are asked to process and incorporate feedback as part of their writing practice using the available resources in P&P and CM writing.

Finally, as our P&P and CM dyads rewrote their texts, they incorporated (successfully or unsuccessfully) into their rewritten texts grammatical and, to a lesser extent, mechanical corrections or merely left the errors uncorrected. This finding shows the correspondence between the types of errors noticed (basically, grammar and spelling) and the types of corrections incorporated. In addition, spelling and grammatical errors (rule-based errors) were easier to understand because of the direct feedback received and thus to incorporate, regardless of the writing environment. Lexical errors, on the other hand, were not so frequent in the drafts and involved wrong lexical choices whose corrections may be more difficult to understand and remember by low proficiency learners. Finally,

the metalinguistic awareness-raising activity focused mostly on the grammatical issues detected in the drafts, which may have driven these learners' attention to grammatical errors.

6. Conclusion

This study aimed to elucidate the effects of writing environment (CM vs P&P) in authentic classrooms on the accuracy of texts rewritten after metalinguistic training and the collaborative processing of teacher WCF. The investigation contributes to the WCF and SLA-oriented L2 writing literature in general by analyzing an underrepresented population in a curricular learning context and by expanding research to include CM writing.

The study showed (i) the effectiveness of Internet resources to improve accuracy when formulating texts; and (ii) the combined effect of unfocused direct feedback, feedback processing, and metalinguistic training to improve the accuracy of low proficiency learners' rewritten texts using authentic curriculum-based tasks conducted in CM and P&P writing. No significant differences were found between groups in feedback processing and the accuracy of rewritten texts because the CM group decided not to use Internet resources beyond Google Docs.

From a methodological viewpoint, we may conclude that processing feedback on a computer does not seem to negatively influence the amount of error noticing and the incorporation of corrections compared to P&P. In addition, giving Internet resources to low proficiency learners does not mean that learners will benefit from them if they are not trained how to use them. We suggest that apart from giving clear guidelines about how to use feedback (see Roca de Larios & Coyle 2022), low proficiency learners also need to be instructed how to use and process this feedback in different writing environments (P&P and CM). In this way, these learners could benefit from the additional resources offered in CM writing (e.g., online dictionaries, grammar checks, corpora) to process feedback collaboratively and revise their texts.

Pedagogically, we suggest that low proficiency learners with little writing experience could write short texts, receive direct corrections and training to notice and explain errors in collaboration (e.g., spot-the-error activities) either in P&P or CM writings so as to increase the likelihood of successfully noticing the errors on which they receive feedback and which they correct in due course. L2 teachers may also want to have their students write their texts on Internet-enabled devices to improve accuracy at the formulation stage of the writing process. Schools that cannot afford this technology could provide students with external resources at this stage, such as students' personal smartphones or tablets. Further studies could investigate if this sustained writing practice with external resources could lead to language learning over time.

This study was conducted with a small sample size of EFL learners, and the use of Internet resources could not be tracked. Future research should collect data on larger samples in other contexts and make use of concurrent data collection instruments such as screencast technology, audio, or video recordings to better capture collaborative writing- and feedback-related processes. Despite these limitations, we consider our study ecologically valid and with genuine implications for SLA research and thus a worthy contribution to the field of writing to learn the language using WCF in P&P and CM environments.

Acknowledgements

This study was funded by the Spanish Ministerio de Ciencia e Innovación (Research grant PID2019-104353GB-I00).

References

Abrams, Z. I. (2019). Collaborative writing and text quality in Google Docs. *Language Learning and Technology, 23,* 22-42. https://doi.org/10125/44681

Adair-Hauck, B., Donato, R., & Cumo-Johanssen (2010). Using a story-based approach to teach grammar. In J. L. Shrum & E. W. Glisan (Eds.), *Teachers' handbook: Contextualized foreign language instruction* (pp. 216-244, 4th ed.). Heinle Cengage Learning.

Alharbi, M. A. (2019). The role of an instructor's asynchronous feedback in promoting students' interaction and text revisions. *International Journal of English Linguistics, 9*(3). https://doi.org/10.5539/ijel.v9n3p23

Bitchener, J. (2019). The intersection between SLA and feedback research. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 85-105). Cambridge University Press. https://doi.org/10.1017/9781108635547.007

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development.* Multilingual Matters. https://doi.org/10.21832/9781783095056

Cánovas Guirao, J., Roca de Larios, J., & Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System*, *52*, 63-77. https://doi.org/10.1016/j.jslw.2018.09.002

Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 173-187). Routledge. https://doi.org/10.4324/9781315165080

Chan, S., Chan, C., Derbie, A. Y., Hui, I., Tan, D., Pang, M., Lau, S., & Fong, K. (2017). Chinese calligraphy writing for augmenting attentional control and working memory of older adults at risk of mild cognitive impairment: A randomized controlled trial. *Journal of Alzheimer's Disease, 58*(3), 735-746. https://doi.org/10.3233/JAD-170024

Cheung, Y. L. (2015). A comparative study of paper-and-pen versus computer-delivered assessment modes on students' writing quality: A Singapore study. *The Asia-Pacific Education Researcher, 25*(1)*, 23-33*. https://doi.org/10.1007/s40299-015-0229-2

Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition, 3,* 451-485. https://doi.org/10.1017/S0272263113000612

Coyle, Y., & Roca de Larios, J. (2020). Exploring young learners 's engagement with models as a written corrective technique in EFL and CLIL settings. *System, 95,* 1-14. https://doi.org/10.1016/j.system.2020.102374

Coyle, Y., Nicolás-Conesa, F., & Cerezo, L. (in press). Overview of methodological procedures in extant research on written corrective feedback processing. In R. M. Manchón & J. Roca de Larios (Eds.), *Research methods in the study of writing processes*. John Benjamins.

DiCamilla, F. J., & *Antón*, M. (1997). Repetition in the collaborative discourse of L2 learners: A Vygotskian perspective. *The Canadian Modern Language Review, 53*, 609-633. https://doi.org/10.3138/cmlr.53.4.609

Elabdali, R. (2021). Are two heads really better than one? A meta-analysis of the L2 learning benefits of collaborative writing. *Journal of Second Language Writing, 52*(2). https://doi.org/10.1016/j.jslw.2020.100788

Ellis, N.C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188. https://doi.org/10.1017/S0272263102002024

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

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

Lee, H. K. (2004). A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing, 9*, 4-26. https://doi.org/10.1016/j.asw.2004.01.001

Leow, R. P. (2020). L2 writing to learn: Theory, research, and a curricular approach. In R. M. Manchón (Ed.), *Writing and language learning: Advancing research agendas* (pp. 95-117). John Benjamins.

Li, J. & Cumming, A. (2001). Word processing and second language writing: A longitudinal case study. *International Journal of English Studies, 1*(2), 127-152.

Manchón, R. M (2014). The distinctive nature of task repetition in writing. Implications for theory, research and pedagogy. *ELIA, 14*, 13-41. https://doi.org/10.12795/elia.2014.i14.02

Manchón, R. M., Nicolás-Conesa, F., Cerezo, L., & Criado, R. (2020). L2 writers' processing of written corrective feedback. In W. Suzuki & N. Storch (Eds.), *Languaging in language learning and teaching: A collection of empirical studies* (pp. 241-263). John Benjamins. https://doi.org/10.1075/lllt.55.11man

Nicolás-Conesa, F., Manchón, R.M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts: Looking into feedback for accuracy and feedback for acquisition. *Modern Language Journal*, *103*(4), 848-873. https://doi.org/10.1111/modl.12592

Owston, R.D., Murphy, S., & Wideman, H. H. (1992). The effects of word processing on students' writing quality and revision strategies. *Research in the Teaching of English*, 26(3), 249-276.

Pearson, W. S. (2022). A typology of the characteristics of teachers' written feedback comments on second language writing. *Cogent Education*, *9*(1). https://doi.org/10.1080/2331186X.2021.2024937

Pennington, M. C. (1996). Writing the natural way: On computer. *Computer Assisted Language Learning, 9*(2-3), 125-142. https://doi.org/10.1080/0958822960090205

Roca de Larios, J., & Coyle, Y. (2022). Learners' engagement with WCF in individual and collaborative L2 writing conditions. In R. M. Manchón & C. Polio (Eds.), *The Routledge handbook of second language acquisition and writing* (pp. 81-93). Routledge.

Saeed, M. A., & Al Qunayeer, H. S. (2020). Exploring teacher interactive e-feedback on students' writing through Google Docs: Factors promoting interactivity and potential for learning. *The Language Learning Journal, 50*(3), *360-377*. https://doi.org/10.1080/09571736.2020.1786711

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11(2)*, 129-158. https://doi.org/10.1093/APPLIN/11.2.129

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge University Press. https://doi.org/10.1017/CBO9781139524780.003

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters. https://doi.org/10.21832/9781847699954

Storch, N. (2018). Research timeline: Collaborative writing. *Language Teaching, 52*(1), 40-59. https://doi.org/10.1017/s0261444818000320

Storch, N. (2022). Theoretical perspectives on L2 writing and language learning in collaborative writing and the collaborative processing of WCF. In R. M. Manchón & C. Polio (Eds.), *The Routledge handbook of second language acquisition and writing* (pp. 81-93). Routledge. https://doi.org/10.4324/9780429199691-4

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, *32*(2), 303-334. https://doi.org/10.1017/S0272263109990532

Suzuki, W. (2012). Written languaging, direct correction, and second language writing revision. *Language Learning, 62*(4), 1110-1133. https://doi.org/10.1111/j.1467-9922.2012.00720.x

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge University Press.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). Continuum.

Vasylets, L., Mellado, M. D., & Plonsky, L. (2022). The role of cognitive individual differences in digital versus pen-and-paper writing. *Studies in Second Language Learning and Teaching*, *12*(4), 721-743. https://doi.org/10.14746/ssllt.2022.12.4.9

Watanabe, Y. & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*(2), 121-142. https://doi.org/10.1177/136216880607074599

Wigglesworth, G., & Storch, N. (2012a). What role for collaboration in writing and writing feedback? *Journal of Second Language Writing*, *21*, 364-374. https://10.1016/j.jslw.2012.09.005

Wigglesworth, G., & Storch, N. (2012b). Feedback and writing development through collaboration: A socio-cultural approach. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 69-99). De Gruyter. https://doi.org/10.1515/9781934078303.69

Zhang, M., & Plonsky, L. (2020). Collaborative writing in face-to-face settings: A substantive and methodological review. *Journal of Second Language Writing*, *49*. 100753. https://doi.org/10.1016/j.jslw.2020.100753

Zhang, M., Gibbons, J., & Li, M. (2021). Computer-mediated collaborative writing in L2 classrooms: A systematic review. *Journal of Second Language Writing*, 54, 1-20. https://doi.org/10.1016/j.jslw.2021.100854

Zhi, M., & Huang, B. (2021). Investigating the authenticity of computer- and paper-based ESL writing tests. *Assessing Writing, 50*, 1-12. https://doi.org/10.1016/j.asw.2021.100548

APPENDIX A

Writing prompts

1. Pre-test: *Write a 150-word paragraph in English describing a) the house, b) the town/city, or c) the beach/mountain area where one of you spend your summer vacation.*
2. Draft: *Write a 150-word paragraph in English describing your bedroom (Student 1's bedroom or Student 2's bedroom only).*
3. Rewriting: *Rewrite your paragraph having in mind the corrections given to you about your grammatical, lexical, and mechanical errors, and the indications provided about the content and the organization of ideas in your paragraph to improve both aspects in this final version.*

APPENDIX B

Instruction 2: Metalinguistic awareness-raising activity

STEP 1: *Whole-class spot-the-error*

Instructions: *Let's identify, correct, and explain together the errors contained in the following sentences.*

Examples:

I have a bed blue ... ➔ I have a blue bed ...

She like eating ... ➔ She likes eating ...

STEP 2: *Collaborative spot-the-error*

Instructions: *With your partner, read the following sentences. Correct and explain the errors contained in them.*

Example:

I have a new mobile phone pink.
CORRECTION: I have a new pink mobile phone.
EXPLANATION: Adjectives are placed before nouns.

STEP 3: *Individual (homework) spot-the-error*

Instructions: *Read the following sentences. Correct and explain the errors they contain.*

Example:

I like play videogames online.
CORRECTION: I play videogames online.
EXPLANATION: 'like' is unnecessary.

APPENDIX C

Feedback processing: Error logs

Step 1: *Put your pencils down! (15 min). Here is the text you wrote together, with your teacher's corrections. Read it aloud, explaining the errors and their corrections to each other. The teacher's corrections are included in red. Attention: Do not write anything at this stage. You must only talk about your errors and the corrections.*

Step 2: *Now, write! (30-35 min). After you have identified your errors and the teacher's corrections and you have talked about them, you must now dictate errors and corrections to each other. Each of you must copy all the errors and corrections in your own worksheet. At the end of the task, you must both have written the same information in your respective worksheets.*

Example:

Error #	Error	Correction
1	of my house	of the house
2	is very comfortable	it is very comfortable
3	I love spent time	I love spending time
4	is like my zone	It is like my zone
5	purple harts	Purple hearts
6	some photos in the walls	some photos on the walls
7	my friends … etc.	my friends …
8	in the tv	on the TV
9	I wach some films	I watch some films
10	I have a big table	I have a big desk
11	there I do my homework	where I do my homework
12	a wardrove	a wardrobe
13	my favorite,	my favorite piece of furniture
14	is like sleep	It's like sleeping
15	I have a big window	There is a big window
16	on the table	on the desk
17	Is not perfect	It is not perfect
18	But is my comfortable zone	But it is my comfortable zone
19	My favorite part of my house	My favorite part of the house

APPENDIX D

Error coding scheme (abridged)

- GRAMMAR
  - Tense
  - Verb Form
  - Word Form
  - Plural
  - Preposition
  - Pronoun
  - Conjunction
  - Discourse Marker
  - Word Order
  - Determiner
  - Sentence Structure
- LEXIS
  - Word
- MECHANICS
  - Spelling
  - Punctuation