Technology Assisted Language Education TALE

Volume 3. Issue 1. March 2025. Pages 133 to 155.

https://doi.org/10.22126/tale.2025.11457.1079

Document Type: Research Paper

Computer-Mediated L2 Collaborative Writing: L1 vs. L2 Interaction Effect on Complexity, Accuracy, Fluency (CAF) and Overall Quality of EFL Learners' Texts

Sharareh Kafian[1], Rajab Esfandiari[2]

[1]MA in English Language Teaching, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran. sharareh_kafian@yahoo.com

[2]Associate Professor of English Language Teaching, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran. esfandiari@hum.ikiu.ac.ir

Received: November 30, 2024**; Accepted:** March 29, 2025

Keywords: Accuracy, Complexity, Computer-Mediated Collaborative Writing, Fluency, Google Docs

Abstract

Despite research studies regarding collaborative writing (CW), the effect of using L1 as the medium of collaboration has been underexplored in computer-mediated L2 CW. This study investigated the effect of the language that the learners employed for collaboration (L1 vs. L2) on their L2 texts and examined whether learners participating in online CW using Google Docs produced better subsequent individual L2 texts. Participants consisted of 45 Iranian upper-intermediate English-as-a-foreign-language (EFL) male and female learners, with Farsi as their mother tongue and their ages ranging from 19 to 24. They were divided into three groups (two experimental groups and one control group). The essays were analyzed in terms of complexity, accuracy, fluency (CAF), and overall quality of the texts to see which language (L1 or L2) led to superior L2 texts. The researchers employed Mann-Whitney U, Kruskal-Wallis, and multivariate analysis of variance (MANOVA) to analyze the test data. The results showed that collaboratively written L2 texts were superior in terms of accuracy and overall quality compared to those generated by the control group. Furthermore, the L1 group performed better regarding complexity, whereas the L2 interaction group was superior in terms of fluency and overall quality of the texts. Based on the findings, the way collaboration is done may play a more important role than the language utilized for collaboration. The findings promise implications for the use of collaborative-based processes to contribute to EFL learners' quality writing.

*Corresponding Author: Rajab Esfandiari

Email: esfandiari@hum.ikiu.ac.ir

Introduction

Considering the undeniable usefulness of CW, any factor that may add to or reduce from the efficacy of this practice is worth investigating. One such factor over which controversies prevail is the use of L1. On the one hand, instructors are usually concerned with whether to allow learners to interact in their L1 while fulfilling L2 CW tasks may decrease their learning opportunities (Zhang, 2021), as the use of L1 might not expose language learners to as much L2 as needed (Swain & Lapkin, 2013). On the other hand, research in this respect has found many advantages that L1 can bring about. For example, Anton and Dicamilla (1999) maintained that the L1-mediated construction of collective scaffolding, which led to retrieving complex linguistic forms, promoted private speech and facilitated maintaining intersubjectivity on the task.

Given the significance attached to collaboration in prior research (e.g, Davison, 2025; Hsu, 2025; Ebadijalal, & Moradkhani, 2025;), insufficient attention to the use of CW processes (whether jointly produced texts or subsequent individually produced writing) may make it impossible to evaluate the amount of learning that takes place during language learners' collaborative experience. Additionally, the findings of most recent research (e.g., Cao et al., 2025) have revealed that not addressing L1/L2 collaboration is most likely to lead to problems in L2 writing, such as not noticing alternative ways of expressing ideas, poor argumentation and weak idea development, reduced motivation and enjoyment in improving writing, poor quality writing (i.e., less complex grammatical structures and inaccurate sentences), among other things. Lastly, as Alsahil (2025) noted, when language learners are engaged in computer-mediated L2 CW processes, the skills and knowledge they acquire are readily transferred to subsequent online texts that they compose independently.

As far as the L1/L2 interaction effect in L2 CW is concerned, two research gaps can be identified. The first gap concerns the use of L1/L2 in naturalistic or controlled settings. The majority of previous research on the use of L1 in CW, as reviewed by Zhang and Li (2023), has been primarily observational studies, in which the amount of L1 use has been measured, or the functions of L1 use have been described, and the language used for interaction has not been controlled. As Zhang (2021) asserted, "this practice renders it difficult for researchers to experimentally investigate how L1 and L2 use may influence different aspects" (p. 341) of L2 CW, such as complexity, accuracy, and fluency (CAF). The second gap relates to the modality in which L1 has been used in collaborative writing. Prior research in L2 CW, as systematically reported in Zhang and Plonsky (2020), Zhang et al. (2021), and Zhang and Li (2023), has primarily been limited to the functions of L1 peer interaction and suggests that the role of L1 in computer-mediated CW on L2 learners' written production such as CAF has comparatively been under-explored. Hence, the following research questions were formulated to achieve the goals of the present study:

Volume 3. Issue 1. March 2025. Pages 133 to 155.

What is the effect of L1 vs. L2 interaction on the CAF of collaboratively produced L2 texts using Google Docs?

What is the effect of L1 vs. L2 interaction on the overall quality of collaboratively produced L2 texts using Google Docs?

What is the effect of L1 vs. L2 interaction on the CAF of subsequent, individually produced L2 texts using Google Docs?

What is the effect of L1 vs. L2 interaction on the overall quality of subsequent, individually produced L2 texts using Google Docs?

Literature review

L2 CW

CW is an instructional activity which is defined by Storch (2011) as "the joint production of a text by two or more writers" (p. 275). Composition scholars have argued that characteristics such as having a specific audience in mind while writing indicate that all kinds of writing may more or less be considered collaborative (Alwahoub et al., 2020; Esfandiari, & Moein; Zhang et al., 2021). However, researchers have emphasized some distinguishing features of CW, all of which have to be present in order to call a practice CW. Ede and Lunsford (1990), for instance, associated CW with three main characteristics of actual interaction among collaborators while composing a text: joint decision-making, responsibility for the produced text, and production of a single written text. As Storch (2013) noted, CW is not applied merely to sharing ideas, and collaborators are involved in all stages of composing, including planning, reaching agreement on form and meaning choices, revising, and editing the text to the point that the final version of the text is created.

Incorporating CW tasks not only fosters learner autonomy but also enhances engagement and peer assessment among L2 learners. Asadi and Taheri (2024), for example, underscore the transformative potential of a multifaceted approach to collaborative writing, particularly in enhancing peer assessment and engagement among L2 learners. Their research illustrates that integrating structured instruction, peer feedback, and advanced technology not only fosters a more interactive and supportive learning environment but also significantly improves students' writing skills. By promoting collaboration and reducing reliance on instructors, this approach empowers learners to take ownership of their educational journey.

Collaboration among peers is claimed to be one of the methods that advance language learners' autonomy and lowers their dependence on instructors (Jacobs & Tan, 2015), thereby making them more responsible in their development process. Furthermore, when learners participate in CW tasks, they engage in collaborative dialogue, a type of languaging (Swain & Watanabe, 2013), defined as "dialogue in which speakers are

engaged in problem solving and knowledge building" (Swain, 2000, p. 102), which means learners are provided with well-timed assistance from their peers in case they experience difficulty doing the task. An additional feature of scaffolding realized in collaborative activities is the provision of assistance that is adjusted based on learners' developmental level. In an English-as-a-second-language (ESL) CW setting, Brooks and Swain (2009) demonstrated that in the posttest stage, participants were mostly able to resolve language problems which they had discussed earlier with their peers rather than those problems for which they had received feedback from the researcher. Overall, since CW tasks help actualize the characteristics of the scaffolding technique, they are worth being employed and investigated in L2 learning and teaching settings as well.

Google Docs in L2 CW

Google has offered various tools to be incorporated in educational settings, such as Gmail, Google Docs, and Google+, among others (Chiablaem, 2021), to enrich teaching and learning practices. Functioning as an online word-processing tool, Google Docs makes it possible for its users to create, edit, and share cloud-based documents with each other (Suwantarathip & Wichadee, 2014). Considering the fact that Google Docs can be accessed by anyone with a computer and internet connection or even with a smartphone using the Google Docs app, Albedaiwi (2022) referred to Google Docs as a suitable tool affording distant pair or group collaborations. According to Kitjaroonchai and Suppasetseree (2021), another affordance of Google Docs is that it can be used in both synchronous and asynchronous modes, meaning that collaborators can make their contribution or edit each other's contributions either in real time or at any time they find convenient. Zioga and Bikos (2020) stated that rather than combining all the work done by different individuals to prepare a joint final document, Google Docs allows collaborators to work simultaneously on one document in a way that while one member is adding something or monitoring the changes applied by others, another member is making changes to the same document. Zioga and Bikos concluded that CW tasks and collective content production have been promoted by this collaborative online word processor.

Due to the useful properties of Google Docs, researchers have been encouraged to investigate its potential merits in regard to teaching and learning writing skills. Numerous such merits have been reported by them, including the advantageous role of Google-Docs-mediated interactive feedback in engaging students in extensive negotiations regarding their writing problems and revision of their texts (Saeed & Al Qunayeer, 2020); increasing students' confidence as well as awareness of writing process and enhancing their individual posttest performance (Suwantarathip & Wichadee, 2014); boosting the efficiency of CW, leading to the production of longer texts, and speeding up the writing process (Apple et al., 2011); enabling teachers to come up with more creative tasks like sharing a text with a lot of mistakes and asking students to correct it

Volume 3. Issue 1. March 2025. Pages 133 to 155.

collaboratively (Zioga & Bikos, 2020); providing teachers with revision histories which act as a way of formative assessment of group members' performance (Zheng et al., 2017); and establishing in learners a more positive attitude towards learning while also increasing their motivation and self-efficacy (Liu & Lan, 2016).

Sociocultural Theory of Mind

Having reviewed previous face-to-face and computer-mediated L2 CW studies, Zhang and Plonsky (2020) and Zhang et al. (2021) concluded that the sociocultural theory of mind has been the most commonly drawn-on theory in L2 CW —the theoretical model the researchers used in this study. Sociocultural theory, originated from the works of the Russian psychologist Lev Vygotsky, emphasizes the defining impact that social and cultural factors have on reorganizing biologically endowed capacities of human beings (such as memory and attention) into a psychological system that is peculiar to human beings (Lantolf, 2006).

As far as mediation is concerned, an important question that L2 researchers have tried to answer is whether adult L2 learners establish the ability to use L2, rather than their L1, to mediate their thinking. Ushakova (1994) stated that even if human beings gain the capability to utilize L2 in their social communications, they still rely on their L1 to regulate their mental activities when facing a problem. Some early studies (e.g., McCafferty, 1994) indicated that ESL learners could indeed mediate and organize their thinking using L2 in such tasks as picture story narration, fairy tale recalling, or explaining the content of an expository text. However, these studies were carried out in an immersion L2 setting rather than a foreign language setting (Lantolf, 2006). The research done by Centeno-Cortés and Jimémez-Jiménez (2004) shed some light on this issue. Centeno-Cortés and Jimémez-Jiménez required participants, who were L1 and L2 speakers of English and Spanish, to solve a number of math, logic, or spatial problems that were cognitively challenging. Analyzing the private speech of participants, Centeno-Cortés and Jimémez-Jiménez noted that although intermediate students read the instructions in L2, almost all of them resorted to their L1 to solve the problem. While some advanced L2 learners managed to continue regulating their cognition using L2, they were not successful at solving the problem until they switched their private speech to their L1. It may, therefore, be the case that regardless of their proficiency level, L2 learners rely on their L1 to regulate their cognitive functions when encountering a challenging problem. Consequently, in the present study, the mediating role of L1 and L2 in CW groups is compared to find which symbolic tool tends to further facilitate cognitive mediation and results in better outcomes. As it is clear from the above discussion, sociocultural theory highlights the crucial role of culturally introduced tools, their mediating effect on cognitive functions, and the fact that such tools are constantly being developed. Accordingly, this study also inspects how Google Docs, as a recent technological tool, mediates learning L2 writing skills.

The zone of proximal development (ZPD) and scaffolding are other dimensions of this theoretical framework. Based on what was explained with regard to regulation in the foregoing paragraph, human activities, including higher-order mental activities such as solving a problem or learning a language, are first realized between people at a social level by other-regulation before being internalized as knowledge to be repeatedly utilized by individuals (Bikowski & Vithanage, 2016). Therefore, when individuals do things under the guidance of more capable people, they exhibit a level of performance which will not be attainable should they work solitarily. This is what Vygotsky (1978) referred to as the ZPD, which he defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). This guidance or assistance from others that individuals are equipped with in the process of collaboration and within their ZPD has been referred to as scaffolding by various scholars (e.g., Doo et al., 2020). Because the present study requires peers to compose joint texts in groups of three, it provides them with the opportunity to engage in collective scaffolding and to learn from one another.

The Role of L1 in Language Learning and L2 CW

The use of L1 in second language learning has been hotly debated and tends to be contentious, and, traditionally, language teachers usually harbor a negative attitude towards using it in L2 classrooms. However, the findings of more recent research studies (e.g., those from a sociocultural theory, Poehner & Lantolf, 2024) have shown the beneficial role of using L1 for language learning. Zhang (2021), for example, has succinctly summarized the research findings and concludes that when students share an L1, "it may not only help students comprehend the target language and elicit more comprehensible input …, but also provide a means for learners to build their social identity and construct the L2 classrooms as a bilingual space … . (p. 340).

Similarly, past research has shown the facilitative role of L1 in L2 CW. Firstly, using L1 serves as a cognitive and communicative resource for language learners (Michel et al., 2025). When language learners produce an L2 text collaboratively, they use their L1 to discuss complex ideas, plan their writing, and clarify their understanding of the task, and the shared L1 helps them to concentrate on higher-order skills, thinking, and reasoning such as organization and argumentation of co-constructed texts (Fogal, 2024). Language learners can use L2 to brainstorm, negotiate meaning, solve linguistic problems, approach CW more effectively, and understand L2 structures more deeply (Zhang & Li, 2023). Additionally, accumulative research findings (see Stell & Iwashita, 2024) suggest that using L1 helps mediate the construction of collective scaffolding, promote the use of private speech, retrieve complex linguistic forms, evaluate and make meaning of the L2 text.

Volume 3. Issue 1. March 2025. Pages 133 to 155.

CAF in L2 CW

In computer-mediated L2 CW, CAF framework has been frequently used to measure the collaboratively and individually produced written products (see Zhang & Li, 2023, for a full review). CAF is considered to be the critical dimension of L2 writing proficiency and development, and several measures have been proposed to evaluate it. As far as empirical research studies are concerned, the findings tend to be mixed when CAF has been examined in L2 CW. McDonough et al. (2018), for example, compared the collaborative and individual texts that 128 Thai EFL students wrote. The results of the study showed that collaboratively produced texts were more accurate than those individually produced, but the individually produced texts were more complex than collaborative texts. One limitation of this study was that the authors did not analyze fluency. Zhang (2018) investigated the impact of L1 and L2 interaction during collaborative writing (CW) tasks on various aspects of Chinese L2 learners' texts, such as CAF and overall quality of writing. Zhang did not find any differences CAF, but the group using L1 produced longer clauses. Similarly, when Zhang (2021) analyzed the effect of L1 and L2 interaction among Chinese L2 students' texts in terms of lexico-grammatical features, she found the use of L1 significantly facilitated the production of complex grammatical features typical of academic writing in learners' co-constructed texts. Villarreal and Gil-Sarratea (2020) analyzed the argumentative essays collaboratively produced by two groups of 32 Spanish L2 students and found that both groups produced less complex texts and more accurate essays but similar fluent texts. However, when Jiang and Eslami (2022) examined effectiveness of CMCW among Chinese L2 students, they found that their essays did not show any significant changes in accuracy and complexity, but improved fluency was observed.

In Iran, several researchers have analyzed L2 CW and CAF, but because of space limitations we restrict ourselves to studies published after 2020 to summarize recently published articles. Zabihi and Bayan (2020), for instance, investigated the CAF of argumentative essays written independently and in pairs by 306 upper-intermediate Iranian EFL university students and found that the essays produced in pairs showed significant improvement, highlighting the role of collaborative writing in CAF. In another study, Hosseini et al. (2021) examined the effect of computer-mediated collaboration on the writing accuracy of 72 pre-intermediate Iranian EFL students at a language institute and found that the experimental group's collaborative essays were more accurate than those of the control group, but Hosseini did not analyze fluency and complexity, and this is considered one limitation of their study. Danandeh and Mohamadi Zenouzagh (2022) investigated the possible effects of multimedia teaching on the CAF of 60 intermediate Iranian EFL learners' collaboratively and individually produced descriptive essays. The results of their data analyses showed that all the students produced more complex, more accurate, and more fluent descriptive essays, but individually produced essays tended to be slightly more fluent than the collaborative essays. Behrahi et al (2024) compared the

effect of three instructional methods (teacher mediation, learner scaffolding, and conventional instruction), using Google Docs, on 97 Iranian EFL learners at a language institute in Khorramabad. The results of posttest showed that the group receiving learner scaffolding outperformed the other two groups in the CAF of their collaborative essays. Cheraghpour et al. (2024) examined how collaborative online writing using Wikis affected 50 Iranian intermediate EFL learners' writing CAF in three language institutes in Shahrekord and found that the experimental group receiving CM instruction significantly improved the CAF of their co-constructed writing.

Method

Participants

Participants in this study consisted of 45 upper-intermediate young adult male and female Iranian EFL learners who were chosen with their own consent from all upper-intermediate students enrolled in each of the five branches of a private foreign language institute in Tehran, Iran. The learners' mother tongue was Farsi, and they ranged from 19 to 24 years of age (Mean age = 21.5; SD = 4.63). Simple random sampling was utilized to randomly select participants and to assign them randomly to three different groups: two experimental groups and one control group. Those learners who had been exposed to CMCW, or in-depth instructions, on how to write problem-solution essays were excluded from the study.

All the participants were also assured that their identities would not be disclosed in reporting the findings of the study, and that they would be provided with the results of the study in case they asked for them. To obtain their consent and ensure ethical considerations, the first researcher of the present study explained the objectives of the study, the procedures for conducting it, and the benefits they may gain from it, and the implications it may have for language learning. After they had been fully briefed on the study procedures, the participants agreed to participate in the study by voluntarily completing the consent forms and submitting them o the researcher.

Instrumentation

Data collection was done during a 5-week writing course from several sources including two pretests, two co-constructed essays using Google Docs, a final individually written essay, and Hedgcock and Lefkowitz' (1992) FL Composition Profile to rate the quality of compositions.

Firstly, to ensure that all the participants had the same level in both their general English and writing skills, two pretests were administered: (a) the second version of a quick placement test (QPT, available online at https://www.vhs-aschaffenburg.de/fileadmin/vhs-aschaffenburg/PDF/OoxfordTest.pdf) by Oxford University Press and University of Cambridge Local Examinations Syndicate, and (b) an

Volume 3. Issue 1. March 2025. Pages 133 to 155.

L2 composition that all participants were asked to produce to confirm the comparability of their initial L2 writing skills. The QPT, the first pretest, included 60 multiple-choice items measuring vocabulary, grammar, reading comprehension, and listening comprehension. The participants were required to take the test in 30 minutes. According to the scoring procedures of the test, those who score between 40 and 47 out of 60 are considered upper-intermediate, or B2, as was the case for the participants in the present study who obtained scores between 40 and 47. QPT is a reliable and valid proficiency test (for more information, see Allan, 2004), but the reliability of the test was re-estimated in the preset study, using Cronbach's alpha, and it was .91, which is considered high enough. The test measures language learners' general language proficiency, but it does not include a writing module. Therefore, to ensure these participants were also homogenous in their initial writing ability prior to the study, the language learners were asked to write a 250-word composition on a topic similar to those given to them in the treatment sessions. Hedgcock and Lefkowitz's (1992) FL Composition Profile was used to rate their essays in terms of content, organization, grammar, vocabulary, and mechanics. The results of Levene's test ($p > .05$) did not show any significant differences in their writing ability to begin the study with. These two pretests were administered to the participants on the first session of the course before the treatment began.

Secondly, all groups in the two interaction conditions (i.e., L1 and L2) were required to use Google Docs to jointly write two problem-solution essays, the second of which was downloaded from Google Docs for further analysis. The reason for incorporating problem-solution essay type in this study was to stimulate interaction among learners, as Storch (2013) advised instructors to choose tasks that "do not easily lend themselves to a division of labor so that learners collaborate rather than cooperate" (p. 159). Finally, the last source of data collection was the problem-solution essays that learners produced online individually, which were analyzed to investigate individual gains from the two different CW experiences learners had been exposed to. Moreover, FL Composition Profile, which is a writing rating scale designed by Hedgcock and Lefkowitz (1992), was used to assess the quality of essays from different stages. This scale considers five constituents of organization, content, vocabulary, grammar, and mechanics on a 0-100-point scale, where each constituent can be rated with one of these four bands: excellent to very good, good to average, fair to poor, and very poor.

We invited two experts in applied linguistics to rate all the essays the 45 language learners produced in the present study. Through a 2-hour rater training session, the objectives of the study and the study procedures were fully explained to the raters, and they were provided with several sample essays already rated through FL Composition Profile in terms of organization, content, vocabulary, grammar, and mechanics. The raters analyzed these samples very carefully, raising their questions concerning the criteria, the scale descriptors, and the rated essays. Next, they were asked to use the FL Composition Profile to rate two essays the participants had produced in this study to ensure they

followed the rating procedures. While the raters were rating the essays, the first researcher of the study monitored the rating session to ascertain they were following the rating procedures. Having been familiarized with the rating criteria and procedures, the two raters were asked to rate all the essays produced by the participants of the study. To ensure the raters were consistent in their ratings of the essays, the researchers estimated the inter-rater reliability of the ratings, following the Cohen's Kappa measure of agreement, and the Kappa coefficient showed very high agreement between the two raters (κ = .97.6), confirming the consistency of the ratings.

Procedure

The first two sessions of the course were dedicated to teaching learners in the experimental groups the overall structure and appropriate vocabulary choice for a problem-solution essay, as well as providing instructions on how to use Google Docs. In each of the next two 60-minute sessions, students in class 1 collaborated using their L1 to write a 120-150-word problem-solution essay in L2 using Google Docs, whereas the participants in class 2 composed such essays with the same topics interacting in L2. The first collaboratively written essay served the purpose of familiarizing participants with their group members and Google Docs affordances. On the fourth week of the course, all the groups in both classes were asked to write their second essays collaboratively under the same circumstances as the first ones, and their final products were collected to be analyzed. On the fifth and last week, all students in both experimental classes were required to utilize Google Docs to compose individually a problem-solution essay as a posttest. The participants in the control group were also asked to write the same essays in the posttest phase.

Data Analysis

In this study, compositions produced in the pretest stage were rated based on a writing rating scale (FL Composition Profile), and the second collaboratively written essays and subsequent individually composed essays were analyzed based on both quantitative measures and the writing rating scale. The quantitative measures were the CAF scores. Drawing on the work of Storch (2005), complexity was measured as the proportion of clauses to T-units (C/T), and the proportion of dependent clauses to clauses (DC/C); and accuracy was measured as the proportion of error-free clauses to all clauses (EFC/C), and the number of errors per word. Fluency was accounted for in terms of the average number of words, T-units, and clauses per text (Chen, 2019). Since the obtained data for these measures from collaboratively written texts (the data collected to answer the first research question) and subsequent individual posttests (the data collected to answer the third research question) were not normal, the first and third research questions were answered by utilizing non-parametric tests of Mann-Whitney U and Kruskal-Wallis tests.

Volume 3. Issue 1. March 2025. Pages 133 to 155.

Regarding the overall quality of essays (research questions 2 and 4), by drawing on the work of Chen (2019), who argued in favor of FL Composition Profile as enabling raters to consider the text as a whole, the writing rating scale created by Hedgcock and Lefkowitz (1992) was utilized to rate the quality of collaborative and individual essays. It is worth mentioning that the second research question compared the data collected from the experimental groups only, while the fourth research question addressed all three groups since it investigated subsequent individual texts. For both the second and fourth research questions, multivariate analysis of variance (MANOVA) was conducted using SPSS 25.0 to investigate the effect of L1 vs. L2 interaction experience on the quality of the experimental groups' final collaborative essays alongside the three groups' subsequent individual essays. For the analysis of the raters' ratings, as explained earlier in this section, inter-rater reliability was used to confirm the consistency of the ratings.

Results

To answer the first research question, the researchers used Mann-Whitney U, because the assumptions for the t-test were not met. As far as accuracy is concerned, the results of Mann-Whitney U test showed no statistically significant difference between the mean ranks of the two groups for the ratio of clauses to T-units, $U = 81.\ 00$, $z = 1.31$, $p = .189$, but the difference was statistically significant for the ratio of dependent clauses to all clauses, $U = 18.\ 00$, $z = 3.94$, $p = .000$. In terms of complexity, the results indicated no statistically significant difference in terms of the proportion of error-free clauses to all clauses ($U = 76.50$, $z = 1.50$, $p = .13$) and the number of errors per word ($U = 94.50$, $z = -.92$, $p = .35$). In terms of fluency, the difference was not statistically significant for the total number of words ($U = 72.00$, $z = 1.68$, $p > .05$) and the average T-units per text ($U = 90.00$, $z = -.97$, $p > .05$). However, a statistically significant difference was found in the average clauses per text, $U = 22.50$, $z = -3.76$, $p < .05$), suggesting that L2 texts produced by groups who employed L2 to collaborate were more fluent than those generated by L1 interaction groups with regard to their average clauses per text.

We used MANOVA to compare the L1 and L2 interaction groups' mean scores on collaboratively produced content, organization, grammar, vocabulary, and mechanics. We checked the assumptions related to MANOVA such as normality, lack of any significant multivariate outliers, lack of multicollinearity, linearity, homogeneity of variances of groups, and homogeneity of covariance matrices, and all of them were met. The results, following Bonferroni adjustment) indicated a statistically significant difference between L1 and L2 groups' overall mean values on collaboratively produced L2 texts, $F\ (5, 24) = 353.45$, $p < .01$, partial $\eta^2 = .987$, representing a large effect size.

Based on the results of the between-subject effects in Table 1 and the mean values in Figure 1, the L2 interaction group had a significantly higher mean than the L1 interaction group on content, $F\ (1, 28) = 7.44$, $p < .01$, partial $\eta^2 = .210$), organization, $F\ (1, 28) = 11.47$, $p < .01$, partial $\eta^2 = .291$), grammar, $F\ (1, 28) = 193.02$, $p < .01$, partial

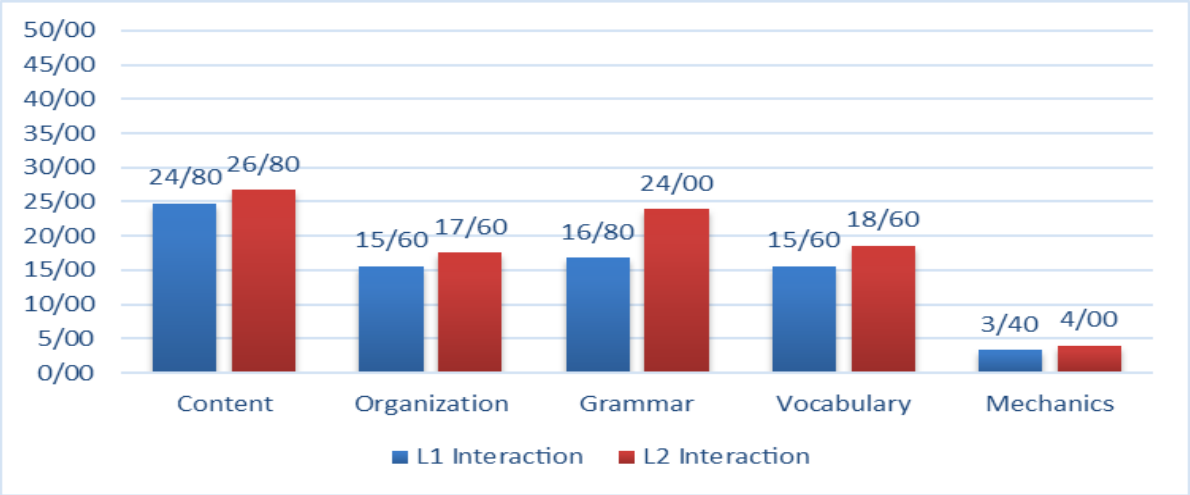
η^2 = .873), vocabulary, F (1, 28) = 30.88, p < .01, partial η^2 = .524). However, the L2 interaction group did not have any significantly higher mean on mechanics than the L1 interaction group, F (1, 28) = 4.84, p > .01).

Table 1
Tests of Between-Subjects Effects for Collaboratively Produced L2 Texts by Group

Source	Dependent Variable	F	Sig.	Partial Eta Squared
Group	Content	7.447	.011	.210
	Organization	11.475	.002	.291
	Grammar	193.021	.000	.873
	Vocabulary	30.882	.000	.524
	Mechanics	4.846	.036	.148

Figure 1 shows the groups' mean values on collaboratively produced texts.

Figure 1
Mean values on Collaboratively Produced Texts



As for the third research question, we ran a series of Kruskal-Wallis tests because the assumptions for the variance analysis were not met. For complexity, the Kruskal-Wallis test did not reveal a statistically significant difference in the ratio of clauses to T-units across the three different groups (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), $\chi^2(2, n = 45) = .30$, p > .05. In ratio of dependent clauses to all clauses (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), χ^2 (2, n = 45) = 4.75, p > .05. As for accuracy, a statistically significant difference in the number of errors per word across the three different groups was found (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), $\chi^2(2, n = 45) = 2.61$, p > .05.

However, no statistically significant difference in proportion of error-free clauses to all clauses was reported (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), $\chi^2(2, n = 45) = 22.97$, p <.05. As for fluency, the Kruskal-Wallis test did not reveal a statistically significant difference in the total number of words across the three different groups (control,

Volume 3. Issue 1. March 2025. Pages 133 to 155.

n = 15, L1 interaction, n = 15, L2 interaction, n = 15), $\chi^2(2, n = 45) = 3.08$, $p > .05$, in average T-unit per text across (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), χ^2 (2, n = 45) = .94, $p > .05$, and in average clauses per text (control, n = 15, L1 interaction, n = 15, L2 interaction, n = 15), χ^2 (2, n = 45) = 1.67, $p > .05$.

The results of MANOVA, F (10, 78) = 20.30, $p < .01$, *partial* η^2 = .422, representing a moderate effect size) showed a statistically significant difference between L1, L2, and control groups' overall means on individually produced texts. Results of the between-subject effects (Table 2) and Scheffe posthoc comparisons (Table 3) revealed the following points. Statistically significant differences between the three groups' mean scores on individually produced content, F (2, 42) = 148.25, $p < .01$, *partial* η^2 = .876 representing a large effect size), revealed differences between the L1 interaction group and the control group, the L2 interaction group and the control group, but not between L2 interaction and L1 interaction.

Table 2

Tests of Between-Subjects Effects for Individually Produced Texts by Group

Dependent Variable	Type III Sum of Squares	Mean Square	F	Sig.	Partial Eta Squared
Content	1108.844	554.422	148.254	.000	.876
Organization	156.311	78.156	35.680	.000	.629
Grammar	461.378	230.689	36.297	.000	.633
Vocabulary	193.644	96.822	29.298	.000	.582
Mechanics	.933	.467	.808	.453	.037

Statistically significant differences were found on individually produced organization, $F(2, 42) = 35.68$, $p < .01$, *partial* η^2 = .629, representing a moderate effect size). The results of post-hoc Scheffe test indicated that the differences were between L1 interaction group and the control group, the L2 interaction group and the control group and the L2 interaction group and the L1 interaction group. There were statistically significant differences on individually produced grammar, F (2, 42) = 33.29, $p < .01$, *partial* η^2 = .633, representing a moderate effect size), with the differences lying between the L1 interaction group and the control group, the L2 interaction group and the control group, and the L2 interaction group and the L1 interaction group. Statistically significant differences were reported on individually produced vocabulary, F (2, 42) = 29.29, $p < .01$, *partial* η^2 = .582, representing a moderate effect size), and the differences were between L1 interaction group and the control group, the L2 interaction group and the control group, and the L2 interaction group and the L1 interaction group. Mechanics was the only variable on which the results did not reach statistical significance.

Table 3

Scheffe Post-Hoc Comparisons Tests for Individually Produced Texts

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Content	L1	Control	9.80*	.706	.000	8.01	11.59
	L2	L1	1.33	.706	.181	-.46	3.13
		Control	11.13*	.706	.000	9.34	12.93

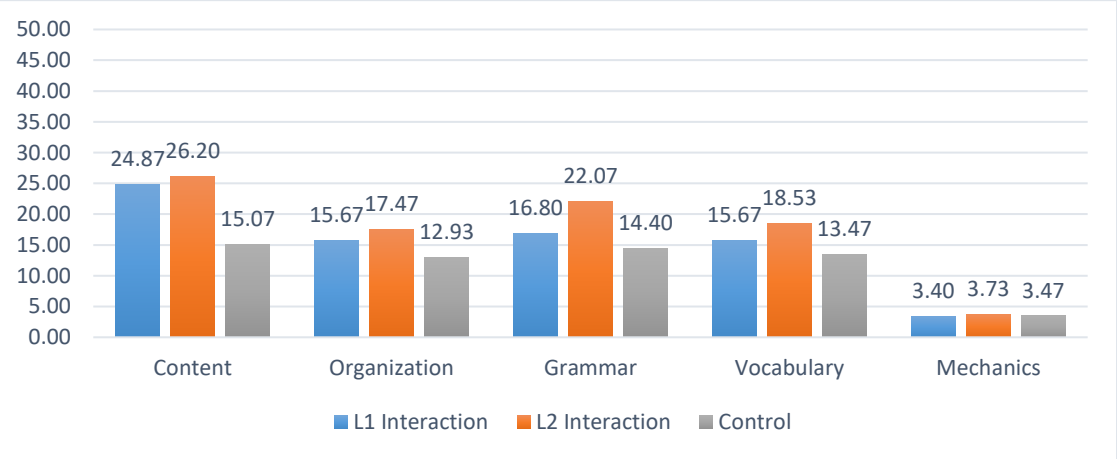
Organization	L1	Control	2.73*	.540	.000 1.36	4.10
	L2	L1	1.80*	.540	.007 .43	3.17
		Control	4.53*	.540	.000 3.16	5.90
Grammar	L1	Control	2.40*	.921	.043 .06	4.74
	L2	L1	5.27*	.921	.000 2.93	7.60
		Control	7.67*	.921	.000 5.33	10.00
Vocabulary	L1	Control	2.20*	.664	.008 .52	3.88
	L2	L1	2.87*	.664	.000 1.18	4.55
		Control	5.07*	.664	.000 3.38	6.75
Mechanics	L2	L1	.33	.278	.492 -.37	1.04
		Control	.27	.278	.633 -.44	.97
	Control	L1	.07	.278	.972 -.64	.77

*. The mean difference is significant at the .05 level.

Figure 2 shows the three groups' mean scores on individually produced texts.

Figure 2

Mean Values on Individually Produced Texts



Discussion

This study examined the existence of any differences in learners' final collaboratively produced L2 texts as well as their subsequent individually produced L2 texts when one group is asked to interact in their L1 (Farsi) and the other in L2 (English) using Google Docs. The finding that L1 interaction resulted in having a greater ratio of dependent clauses to all clauses suggests that the essays included more complex written products at the clausal level. The inclusion of more complex clauses typically means the mean length of the clause is higher in the target texts. The mean length of a clause is a useful indicator of complexity in a variety of writing-related studies (e.g., Norris & Ortega, 2009). For example, Lu (2011) employed 14 metrics pertaining to syntax to study texts composed by L2 students of different levels. When there was an increase in their proficiency levels, an increase could be seen in the mean length of clauses, which implied that the longer the clauses, the higher their complexity.

Volume 3. Issue 1. March 2025. Pages 133 to 155.

Such features as more complex and longer clauses have been regarded as signs of academic writing in other studies (e.g., Staples et al., 2016). In light of this, the finding that students produced lengthier clauses in the texts appears to suggest that the language they used for their essays was more closely aligned with academic registers. Put simply, engaging in L1 interaction enables students to package knowledge using language, a crucial skill for writing in academic settings. This finding confirms that of Zhang (2018), who found that when Chinese L2 learners used their L1 to co-construct texts, the texts tended to be more syntactically complex, including longer clauses. However, the finding is not in line with some earlier research studies (e.g., Fernandez Dobao, 2012; Wigglesworth & Storch, 2009). For example, Wigglesworth and Storch examined 96 L2 texts produced collaboratively and individually by international students at an Australian university and reported no difference in the complexity of the texts. Similarly, Fernandez Dobao (2012) analyzed the L2 texts Spanish L2 students produced individually and in groups of two and four and found no clear differences in the complexity of individual and collaborative texts.

The observation that only clause-related complexity showed a significant difference, rather than the total complexity by subordination (i.e., the number of clauses per T-unit), may be attributed to the different focuses of this measure. According Yoon (2017), writers can extend information by using clausal embeddings, which are associated with clauses per T-unit. Phrasal embeddings, however, are used to compress information and are present in features like nominalizations and prepositional phrases that are more common in L1 use condition. These phrasal embeddings are less likely to meaningfully contribute to T-unit-based measurements than clausal embeddings. Therefore, findings may indicate that L1 interaction permits students to utilize more features for information compression rather than information elaboration. Further investigation is needed to determine the precise causes of this trend in clausal and phrasal embeddings by examining how students retrieve linguistic devices in the dyadic interactions during CW activities.

Regarding accuracy, although mean ranks of L2 interaction groups were higher, the differences of accuracy measures between the two groups were not statistically significant. This finding is in line, to some extent, with that of Zhang (2018), who suggested that L1 interaction has no prominent influence on the written accuracy in students' collaborative texts. The finding is also consistent with that of Jianga and Eslami (2022), who investigated the L2 essays individually produced by Chinese EFL learners at a Chinese research university and found that the accuracy of the L2 individually produced texts relatively remained intact. By contrast, our finding does not confirm that of Sanchooli et al. (2022) who reported that, when Iranian EFL learners received synchronous and asynchronous computer-mediated collaborative practice, they produced L2 texts which were more accurate, as measured by a posttest. Considering the fact that the variation in this study is not statistically significant, more investigation is required to

determine the possible beneficial effects of using learners' mother tongue on linguistic accuracy.

In terms of fluency, especially in the total number of words and the average T-units per text, L1 and L2 users showed the same results. Nevertheless, the students who received L2 produced more fluent texts than those receiving L1 in the average clauses per text. Although less research has been done on writing fluency, Namaziandost et al. (2020) studied the effect of cooperative approaches on EFL learners' fluency development. Their findings revealed a positive correlation between learning cooperatively and language learners' fluency. The findings of this study confirm those of Namaziandost et al. (2020) regarding the higher use of average clauses per text in the L2 group. Our finding partly supports that of Jianga and Eslami (2022), who found the most gain in fluency among Chinese L2 learners. Furthermore, Wang et al. (2025) reported that the Chinese EFL learners who practiced collaborative writing tended to produce more fluent written products. On the other hand, our findings regarding fluency contradict the findings obtained by Zhang (2019), who found no significant differences which partly accounts for the similarity in the total number of words and the average T-units per text in this study.

The study's results on text quality (particularly in terms of content, structure, grammar, and vocabulary) are consistent with those of Swain and Lapkin (2013), who discovered a negative relationship between text quality and the amount of L1 use. Two reasons may account for the findings' similarity: (a) the tasks—writing argumentative essays—were the same; (b) the participants came from similar backgrounds (EFL). Nonetheless, there are conflicting results. For example, Zhang (2018) found no differences between the groups using L1 or L2. The task type used in this study—writing argumentative essays—may have something to do with this variation in text quality. In meaning-focused writing assignments like producing argumentative essays, pairs are most likely to pay close attention to content because it is essential to such tasks (assuming that the two languages being used for interaction have the same writing techniques). Additionally, the closeness in the quality of mechanics between L1 and L2 use circumstances could have been attributed to students' preexisting knowledge. The results somewhat support the previous findings of Elola and Oskoz (2010), who found no statistically significant variations in accuracy metrics between writing done individually and in collaboration.

Zhang (2018) found no indication of improved fluency with L1 use in the collaborative writing process, which is consistent with the nonsignificant impacts of collaborative and subsequent individually generated L2 texts regarding fluency in the three groups in this study. These results could be explained by the same exposure to group writing exercises: Zhang's research participants engaged in either an L1 or L2 group writing exercise. However, Mozaffari (2017) found that collaborative writing can improve fluency under specific pairing conditions (teacher-selected pairs), which is in

Volume 3. Issue 1. March 2025. Pages 133 to 155.

contrast to the nonsignificant effects of collaborative writing on improving fluency in collaborative and subsequent individually produced L2 texts in the current study. Overall, this study builds on previous research by painting a more nuanced picture of the potential contribution of collaborative writing to the level of fluency in L2 texts that are generated jointly as well as later independently.

Lastly, the ensuing individual texts' complexity did not differ significantly among the three groups. The finding might have been influenced by the participants' English proficiency and the learning objectives for this particular course. Thus, these findings along with those of other research (e.g., Mozaffari, 2017) suggest that high complexity is not necessarily a product of CW. These results imply that, in order to increase the texts' complexity, collaborative writing must be used wisely and sparingly. Zhang (2018) discovered, for instance, that pairs who were able to communicate well in their L1 tended to write writings that were more sophisticated.

Improved grammar and vocabulary persisted in both the immediate and subsequent L2 writings. This finding confirms that of Chen's (2019) study. This is one relatively surprising finding when responding to the fourth research question about the quality of subsequent independently produced L2 texts. Only the collaborative groups, though, demonstrated better organization. Part of the reason for this outcome is that students paid close attention to using proper grammar and vocabulary when writing. The results of writing quality in immediate and subsequent texts should also be compared to Shehadeh's (2011) qualitative assessments, which indicated that collaboration had a substantial impact on content, but not grammar. Two factors may have contributed to the differences: The variations in allocated responsibilities and the participants' varying levels of language skill. We are unable to make any assumptions about the impact of cooperation on writing quality for texts that are created collaboratively or independently without more research because there are few studies that qualitatively assess this effect.

The findings of the study can be partly interpreted in light of the sociocultural theory explained in the literature review. As the findings have shown, EFL learners can use their L1 to mediate cognitive functioning and argumentation to produce more complex argumentative essays, reflecting higher-order thinking and reasoning involved in the argumentative process. Similarly, they can use their L2 to regulate witting processes to produce more fluent essays and manage to succeed in achieving better overall writing performance. As such, EFL learners' L1 and L2 may act as semiotic tools to mediate their cognition, helping them to solve a complex problem, such as the use of more sophisticated grammatical structures to produce a problem-solution essay as well the use of various linguistic elements such as content, vocabulary, and organization to improve witting quality.

Conclusion

CW can be considered a beneficial activity to be incorporated in language classes. Based on the sociocultural theory (Lantolf & Xi, 2023), when students are engaged in collaboration with their peers who might own different resources, they can collectively scaffold one another to improve their overall performance and learn from each other. Because no significant difference was found in measures of accuracy between the L1 and L2 interaction groups, and considering the fact that L1 usage even led to more complex L2 texts, it may be concluded that using L1 as the medium of L2 CW activities might not be as harmful as conceived by some L2 instructors (Azkarai, 2023).

The findings may promise several implications. From a theoretical point of view, the results provide additional support for sociocultural theories on CW instructions. This research contributes to the growing scholarly research on CW and supports the tenet that the CW method, considering online sharing platforms, can be successful in language education. As such, language teachers might be encouraged to use collaboration to enhance the quality of the teaching process, as teachers should always pay attention to incorporating new teaching techniques into their courses and help students to be more active in the classes. From a pedagogical point of view, material developers might also benefit from the impact of CW instructions on EFL learners' writing improvement. Material developers may include several exercises in textbooks that need pair or group work to help teachers use and language learners learn them.

Despite the implications, the present study suffers from two limitations. The first limitation relates to the sampling procedure and research setting in which the participants were selected. We selected the participants from a single language institute, and the findings of the study, as an anonymous reviewer also reminded us, may not be generalizable to other similar contexts. This is most likely to affect the external validity of the study, and the in future, researchers should consider using stratified random sampling to choose participants from as diverse EFL contexts as possible to strengthen validity and confirm generalizability purposes. Another limitation concerns the study procedure. The present study primarily dealt with a limited number of sessions and essay topics, as a result of which the findings may not be generalizable to other similar contexts, settings, and population, as an anonymous reviewer asserted. While this number of sessions and topics may be considered legitimate and acceptable, given the cross-sectional nature of the present study, we recommend that researchers, in the future, follow longitudinal research designs to enable them to cover more essay topics in as a large number of sessions as possible.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest: The authors declare that they have no competing interests.

Volume 3. Issue 1. March 2025. Pages 133 to 155.

References

Albedaiwi, S. A, (2022). Collaborative writing on a digital platform: Measuring gains of EFL learners in Saudi Arabia. *Journal of Language and Linguistic Studies, 18*(Special Issue 1), 760-770.

Allan, D. (2004). *Oxford placement test*. Oxford University Press

Alsahil, A. (2025). Exploring students' perceptions and affordances of Google docs-supported collaborative writing. *Innovation in Language Learning and Teaching, 19*(1), 64-82.

Alwahoub, H. M., Azmi, M. N. L., & Halabieh, M. (2020). Computer-assisted collaborative writing and students' perceptions of Google Docs and Wikis: A review paper. *Language Literacy: Journal of Linguistics, Literature and Language Teaching, 4*(1), 15–27. https://doi.org/10.30743/ll.v4i1.2499

Anton, M., & Dicamilla, F. J. (1999). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Modern Language Journal, 83*(2), 233-247. https://doi.org/10.1111/0026-7902.00018

Apple, K. J., Reis-Bergan, M., Adams, A. H., & Saunders, G. (2011). Online tools to promote student collaboration. Getting connected: Best practices for technology enhanced Teaching and Learning in High Education, 239-252. https://doi.org/10.1093/acprof:osobl/9780199733187.003.0016

Asadi, M., & Taheri, R. (2024). Enhancing peer assessment and engagement in online IELTS writing courses through a teacher's multifaceted approach and AI integration. *Technology Assisted Language Education, 2*(2), 94-117. https://doi.org/10.22126/tale.2024.11083.1058

Azkarai, A. (2023). L1 use and patterns of interaction of young EFL learners in a collaborative writing task. *Language Teaching for Young Learners, 5*(1), 7-29. https://doi.org/10.1075/ltyl.22010.azk

Behrahi, L., Khalaji, H. R., & Rajabi, P. (2024). A mixed methods study of the impact of online teacher mediation and learner scaffolding via Google Docs on EFL learners' writing complexity, accuracy, and fluency. *Technology Assisted Language Education, 2*(3), 1-24.

Bikowski, D., & Vithanage, R. (2016). Effects of web-based Collaborative writing on individual L2 writing development. *Language Learning & Technology, 20*(1), 79–99. https://www.learntechlib.org/p/176108

Brooks, L., & Swain, M. (2009). Languaging in collaborative writing: Creation and response to expertise. In A. Mackey and C. Polio (Eds.). *Multiple perspectives on interaction in SLA* (pp. 58–89). Lawrence Erlbaum.

Cao, Z., Ruegg, R., & Skalicky, S. (2025). Manner and quality of negotiation during L2 collaborative writing. *Applied Linguistics, 46*(1), 35-52.

Cheraghpour, Z., Hosseinpour, N., & Shafiee, S. (2024). Effects of Wiki-mediated collaborative writing on Iranian intermediate EFL learners' written complexity, accuracy, and fluency. *International Journal of Foreign Language Teaching and Research, 12*(48), 79-96.

Chiablaem, P. (2021). Enhancing English communication skills of Thai university students through Google apps for education (GAFE) in a digital era during COVID-19 pandemic. *Shanlax International Journal of Education, 9*(3), 91-98. https://doi.org/10.34293/education.v9i3.3921

Danande, S., & Mohamadi Zenouzagh, Z. (2022). The effect of individual and collaborative writing on fluency, accuracy, and complexity of writing performance of Iranian intermediate learners: Mediating role of multimodal instruction. *ZABANPAZHUHI (Journal of Language Research), 14*(44), 115-144.

Davison, I. (2025). The effects of completing collaborative or independent writing on the development of language use in individual writing. *Journal of Second Language Writing, 65*, Article e101128. https://doi.org/10.1016/j.jslw.2024.101128

Doo, M., Bonk, C. & Heo, H. (2020). A meta-analysis of scaffolding effects in Eckerth, J. (2008). Investigating consciousness-raising tasks: Pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics, 18*(2), 121–145. https://doi.org/10.1111/j.1473-4192.2008.00188.x

Ebadijalal, M., & Moradkhani, S. (2025). Impacts of computer-assisted collaborative writing, collaborative prewriting, and individual writing on EFL learners' performance and motivation. *Computer Assisted Language Learning*, *38*(1-2), 291-315.

Ede, L., & Lunsford, A. (1990). *Singular texts/plural authors*. Southern Illinois University Press.

Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering L2 development and mastery of writing conventions. *Language Learning and Technology*, *14*(3), 51–71. http://llt.msu.edu/vol14num3/elolaoskoz.pdf

Esfandiari, R., & Moein, G. (2015). A corpus-driven food science and technology academic word list. *Issues in Language Teaching*, *4*(2), 157-131.

Fernandez Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*, 40-58.

Fogal, G. G. (2024). Expanding the collaborative writing research framework: A longitudinal analysis of how collaborative and independent writers orient to writing spaces. *Journal of Second Language Writing, 63*, Article e101096. https://doi.org/10.1016/j.jslw.2024.101096

Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language writing, 1*(3), 255–276. https://doi.org/10.1016/1060-3743(92)90006-B

Hosseini, M. S., Bavali, M., & Rezvani, R. (2021). Wiki-based collaborative writing in EFL classrooms: Writing accuracy in focus. *Journal of Language and Translation, 11*(1), 65-75. http://dx.doi.org/10.30495/TTLT.2021.679446

Housen, A., De Clercq, B., Kuiken, F., & Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second Language Research, 35*(1), 3-21. https://doi.org/10.1177/0267658318809765

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In Housen, A., Kuiken, F., & Vedder, I. (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1-20). John Benjamins Publishing Company.

Hsu, H. C. (2025). The effect of collaborative prewriting on L2 collaborative writing production and individual L2 writing development. *International Review of Applied Linguistics in Language Teaching*, 63(1), 499-533.

Jacobs, G. M., & Tan, H. S. (2015). Advancing learner autonomy in EFL via collaborative learning. In Handoyo Puji Widodo (Ed.). *Proceedings from the 7th International Conference on Teaching English as Foreign Language (COTEFL)*. Purwokerto, Indonesia: Faculty of Letters, University of Muhammadiyah.

Jiang, W., & Eslami, Z. R. (2022). Effects of computer-mediated collaborative writing on individual EFL writing performance. *Computer Assisted Language Learning*, *35*(9), 2701-2730.

Kitjaroonchai, N., & Suppasetseree, S. (2021). A case study of ASEAN EFL learners' collaborative writing and small group interaction patterns in Google Docs. *English Language Teaching, 14*(5), 89-108.

Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in second language acquisition*, *28*(1), 67-109. https://doi.org/10.1017/S0272263106060037

Lantolf, J. P. (2011). The sociocultural approach to second language acquisition: Sociocultural theory, second language acquisition, and artificial L2 development. In *Alternative approaches to second language acquisition* (pp. 24-47). Routledge.

Lantolf, J. P., & Xi, J. (2023). Digital Language Learning: A Sociocultural Theory Perspective. *TESOL Quarterly*, *57*(2), 702-715. https://doi.org/10.1002/tesq.3218

Liu, S. H. J., & Lan, Y. J. (2016). Social constructivist approach to web-based EFL learning: Collaboration, motivation, and perception on the use of Google Docs. *Journal of Educational Technology & Society*, *19*(1), 171-186. https://www.jstor.org/stable/jeductechsoci.19.1.171

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL quarterly*, *45*(1), 36-62. https://doi.org/10.5054/tq.2011.240859

McCafferty, S. G. (1994). The use of private speech by adult ESL learners at different levels of proficiency. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 117–134). Ablex.

McDonough, K., De Vleeschauwer, J., & Crawford, W. (2018). Comparing the quality of collaborative writing, collaborative prewriting, and individual texts in a Thai EFL context. *System*, *74*, 109-120. https://doi.org/10.1016/j.system.2018.02.010

Michel, M., Bazhutkina, I., Abel, N., & Strobl, C. (2025). Collaborative writing based on generative AI models: Revision and deliberation processes in German as a foreign language. *Journal of Second Language Writing*, *67*, Artile e101185. https://doi.org/10.1016/j.jslw.2025.101185

Mozaffari, S. H. (2017). Comparing student-selected and teacher assigned pairs on collaborative writing. *Language Teaching Research*, *21*, 496–516. https://doi.org/10.1177/1362168816641703

Namaziandost, E., Homayouni, M., & Rahmani, P. (2020). The impact of cooperative learning approach on the development of EFL learners' speaking fluency. *Cogent Arts & Humanities*, *7*(1).

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Poehner, M. E., & Lantolf, J. P. (2024). *Sociocultural theory and second language developmental education*. Cambridge University Press.

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*, *12*(6), 1-18. https://doi.org/10.1080/09571736.2020.1786711

Sanchooli, A., Okati, F., & Mojavezi, A. (2022). The effect of synchronous and asynchronous computer-mediated communication (CMC) on Iranian EFL learners' writing accuracy and complexity. *Iranian Journal of English for Academic Purposes, 11*(1), 13-20.

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of second language writing, 20*(4), 286-305. DOI: 10.1016/j.jslw.2011.05.010

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), 149-183. https://doi.org/10.1177/0741088316631527

Stell, A., & Iwashita, N. (2024). Enhancing collaboration: Exploring regulated learning strategies in the co-regulatory processes of collaborative L2 writing. *System, 125*, Article e103410. https://doi.org/10.1016/j.system.2024.103410

Storch, N. (2005). Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. *Annual Review of Applied Linguistics, 31*, 275–288. http://dx.doi.org/10.1017/S0267190511000079

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

Suwantarathip, O., & Wichadee, S. (2014). The effects of collaborative writing activity using Google Docs on students' writing abilities. *Turkish Online Journal of Educational Technology, 13*(2), 148-156.

Swain, M. (2000). The output hypothesis and beyond. In J.P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.

Swain, M., & Lapkin, S. (2013). A Vygotskian sociocultural perspective on immersion education: The L1/L2 debate. *Journal of Immersion and Content-based Language Education, 1*(1), 101–129. https://doi.org/10.1075/jicb.1.1.05swa

Swain, M., & Watanabe, Y. (2013). Collaborative dialogue as a source of learning. In C. Chapelle (Ed), *The encyclopedia of applied linguistics* (pp. 289-310). Wiley Blackwell.

Ushakova, T. (1994). Inner speech and second language acquisition: An experimental-theoretical approach. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 135–156). Ablex.

Villarreal, I., & Gil-Sarratea, N. (2020). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research, 24*(6), 874-897. https://doi.org/10.1177/1362168819829017

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Harvard University Press.

Wang, K., Zhang, L. J., & Cooper, M. (2025). Metacognitive instruction for improving the effectiveness of collaborative writing for EFL learners' writing development. *The Asia-Pacific Education Researcher*, *34*(2), 661-673.

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26*(3), 445-466.

Yoon, H. J. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System, 66*, 130-141.

Volume 3. Issue 1. March 2025. Pages 133 to 155.

Zhang, M. (2018). Collaborative writing in the EFL classroom: The effects of L1 and L2 use. *System, 76*, 1-12. https://doi.org/10.1016/j.system.2018.04.009

Zhang, M. (2019). Towards a quantitative model of understanding the dynamics of collaboration in collaborative writing. *Journal of Second Language Writing, 45*, 16–30. https://doi.org/10.1016/j.jslw.2019.04.001

Zhang, M. (2021). Understanding L1 and L2 interaction in collaborative writing: A lexico-grammatical analysis. *Language Teaching Research, 25*(3), 338-359. https://doi.org/10.1177/1362168819859911

Zhang, M., & Li, M. (2023). Collaborative writing in face-to-face and computer-mediated L2 settings: A methodological perspective. In M., Zhang, & M. Li. (Eds.), *L2 collaborative writing in diverse learning contexts* (pp. 53-77). John Benjamins Publishing Company.

Zhang, M., & Plonsky, L. (2020). Collaborative writing in face-to-face settings: A substantive and methodological review. *Journal of Second Language Writing, 49*, Article 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*, Article e100854.

Zheng, B., Yim, S., & Warschauer, M. (2017). Social media in the writing classroom and beyond. *The TESOL encyclopedia of English language teaching*, 1-5. https://doi.org/10.1002/9781118784235.eelt0555

Zioga, C., & Bikos, K. (2020). Collaborative writing using Google docs in primary education: development of argumentative discourse. *Turkish Online Journal of Distance Education, 21*(1), 133-142. https://doi.org/10.17718/tojde.690372