Prewriting in Voice Chat: Exploring the Effects of Collaborative Prewriting on EFL Learners' Performance

VARARIN CHAROENCHAIKORN

Chulalongkorn University Language Institute, Thailand
Author email: cvararin@hotmail.com

Article information	Abstract
	Addressing the lack of studies on L2 online and collaborative prewriting, this study examines L2 collaborative prewriting through voice chat. It explores the effects of an online prewriting condition (individual vs collaborative) on performance in subsequent individual writing as assessed by text length, accuracy, complexity, and analytical rating scales of content, organization, and language. One hundred and twenty-six Thai university students were divided into a control (individual prewriting) group and an experimental (collaborative prewriting) group. The utilised prewriting task was structured, focusing on ideas generation, selection, and organization, and the main task was problem-solution paragraph writing. Independent-sample t-tests revealed that accuracy was comparable across groups; however, the experimental group significantly outperformed the control group on text length, some lexical and syntactic complexity measures, and all rating scores, and the effect size concerning text length was large. Overall, these findings suggest potential benefits of collaboration in voice-chat prewriting.

INTRODUCTION

Nowadays, online language learning is becoming increasing popular. It reduces the constraints of time and space and affords interaction and collaboration, which may facilitate language learning, as learners can interact with other learners or native speakers online. It has also been hypothesized that, owing to the sense of anonymity, some online platforms offer learners a more relaxed environment for interaction compared to the traditional face-to-face mode (Beauvois, 1997), which might lead to increased interaction/ participation (Chun, 1994; Kern, 1995; Sullivan & Pratt, 1996). In the realm of writing, research evidence shows that L2 online collaboration may benefit language learning in many ways, e.g. writing improvement and quality (Abrams, 2019; Bikowski & Vithanage, 2016; Strobl, 2014; Wang, 2015).

Recently, many L2 researchers have paid attention to online collaborative writing, while online collaboration during the L2 prewriting stage has remained under-researched, mirroring the lack of collaborative prewriting (CP) research in the non-computer-assisted language learning context. To the best of my knowledge, previous research on L2 online prewriting has been scarce and limited to text-chat mode. The present research aims to address this issue and

extend the scope of the investigation of L2 online prewriting to voice chat, which will allow language researchers and educators to gain a deeper understanding of online prewriting. In particular, it explores prewriting discussion in voice chat and compares the effects of L2 collaborative and individual online prewriting.

LITERATURE REVIEW

Prewriting

Prewriting is the strategy adopted prior to writing to facilitate it. Prewriting strategies include, for instance, outlining, freewriting, listing, clustering, and prewriting discussion (Kroll, 2002; Weigle, 2014). The benefit of prewriting is that it enables writers to establish their goals, generate and organize content, and design the structure of their writing (Ellis & Yuan, 2004; Flower & Hayes, 1981; Hayes & Nash, 1996). It also allows writers to review relevant lexical items and ideas and focus on different language aspects, thereby possibly producing a better-quality text (Williams, 2005). In essence, the advantages of prewriting are supported by Skehan's (1998) Limited Attentional Capacity Model and Kellogg's (1988, 1990) Overload Hypothesis which assume human attention span is limited. Skehan proposed that, due to this limitation, L2 writers may struggle to attend to all three performance aspects, complexity, accuracy and fluency, simultaneously during tasks, leading to a trade-off effect between form (accuracy and complexity) and content (fluency), and between complexity and accuracy within form. Pre-task planning (i.e. strategic planning) may reduce the cognitive demands during tasks and, in turn, ease this trade-off effect. Similarly, the Overload Hypothesis assumes that pre-task planning frees up writers' working memory, which is limited in capacity, allowing writers to focus more on the translation process (i.e. translating ideas into language) during writing and, hence, resulting in increased writing fluency and higher text quality. This hypothesis is based on the notion that writing is a complex activity entailing multiple processes – formulation, execution, and monitoring, each potentially utilising cognitive resources. However, despite these hypotheses, concerns have also been raised regarding prewriting. For instance, it is thought that planning might reduce writing quality as it potentially makes writing less dynamic and spontaneous by reducing the interaction between different writing processes, such as planning, translating, and monitoring (Kellogg, 1990). In addition, as online planning, or planning during writing, naturally occurs in writing, the effects of prewriting may not be obvious (Johnson et al., 2012).

To date, L2 pre-task planning research has primarily focused on exploring planning before an oral performance, revealing positive findings about performance (e.g. Ellis, 2009; Foster & Skehan, 1996; Levkina & Gilabert, 2012; Yuan & Ellis, 2003). Fewer L2 pre-task planning studies have been conducted on prewriting and most of them have investigated individual prewriting (IP). In the realm of L2 IP, findings show some positive evidence. Many studies have found that L2 IP may lead to increased fluency (Ellis & Yuan, 2004; Ghavamnia et al., 2013; Johnson et al., 2012; Ojima, 2006; Rostamian et al., 2018) and syntactic complexity (Ellis & Yuan, 2004; Ghavamnia et al., 2013; Ojima, 2006; Rostamian et al., 2018), and some have found positive outcomes for content quality (Ong & Zhang, 2013). Taken together, these point to the possible effects of prewriting in preparing learners for the task and reducing cognitive demands during

writing. However, evidence for the benefits of L2 IP regarding accuracy is limited and, due to limited research, more studies in this area are warranted to provide conclusive findings.

Collaboration and L2 prewriting

While pre-task planning research has mostly focused on planning for speaking and, in the case of writing, IP, L2 CP could prove beneficial. The usefulness of collaborative tasks is supported by several theoretical perspectives. According to interactionist theories, negotiation of meaning, output generation, and reception of comprehensible input during interaction are key to L2 development, and interaction affords opportunities for hypothesis testing, receiving feedback, and noticing gaps in one's L2 knowledge (Long, 1996; Schmidt, 1990; Swain, 1995). In a sociocultural view (Vygotsky, 1978), interaction is important for the co-construction of knowledge as it enables peer scaffolding, allowing interlocutors to achieve outcomes they could not achieve by themselves. To date, research on collaboration in L2 writing has shown that, during collaborative writing, learners provide immediate language feedback to their peers, negotiate meanings, and share information about content and language (e.g. Fernández Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009); in addition, a recent meta-analysis by Elabdali (2021) revealed that texts written collaboratively are potentially more accurate than those written individually, and learners' experience with collaborative writing may lead to gains in the post-test individual writing stage. Overall, accumulated evidence shows the potential of collaboration for L2 writing development.

Unlike collaborative writing, however, CP does not entail collaboration throughout the whole writing process, which includes planning, composing, editing, and revising, and the final writing product does not have shared ownership. Similar to IP, CP allows writers to plan both content and language, but with their peers, before performing a writing task individually. According to Liao (2018), planning taking place in the prewriting stage essentially differs from that occurring at the time of writing and, consequently, should be viewed in its own right. Even so, due to its collaborative nature, it is possible that CP offers similar benefits to collaborative writing. Moreover, in certain situations CP may be more advantageous than collaborative writing as it takes less time to implement and enables educators to assess the writing performance of each learner separately while simultaneously deriving benefits from peer interaction (McDonough et al., 2018).

To date, L2 CP studies are quite limited. Research in this area has explored prewriting using both structured (e.g. McDonough et al., 2019; Neumann & McDonough, 2014, 2015; Pospelova, 2021) and unstructured (e.g. Kang & Lee, 2019; McDonough et al., 2018) prewriting tasks in both offline (e.g. Pospelova, 2021; Shin, 2008) and online (e.g. Amiryousefi, 2017; Kessler et al., 2020; Liao, 2018) modes. These studies have examined the effects of L2 CP on individual writing performance through rating scores (McDonough et al., 2018, 2019; Pospelova, 2021; Shi, 1998; Shin, 2008) and complexity, accuracy, and fluency measures (Abrams & Byrd, 2017; Kang & Lee, 2019; McDonough et al., 2018, 2019); writing development over time (McDonough & De Vleeschauwer, 2019); and motivation (Lan et al., 2015). Some have focused on comparing teacher- and student-led prewriting tasks (Amiryousefi, 2017; Shi, 1998) or collaborative meaning- and grammar-focused prewriting (Abrams & Byrd, 2017), while others have explored

the language and strategies learners use during prewriting (Joaquin et al., 2016) or the connection between prewriting and other factors, including L2 proficiency (Pospelova, 2021; Shin, 2008), task complexity (Kang & Lee, 2019), and learners' preference for collaboration (Neumann & McDonough, 2014). In research examining prewriting discussions, researchers have investigated the characteristics of these discussions (McDonough & González, 2020; McDonough & Neumann, 2014; Neumann & McDonough, 2015) and their relationship to text quality (McDonough et al., 2018, 2019; Neumann & McDonough, 2015).

L2 collaborative prewriting and writing performance

In light of the research on writing performance, a study by McDonough et al. (2019) examined prewriting in 57 Thai notice writers of English, utilising a structured prewriting task focusing on content and organization prior to a problem-solution paragraph writing task. The CP group in this study significantly outperformed the IP group on accuracy and rating scores, while there was no impact on syntactic complexity. The positive influence on ratings detected lends support to Neumann and McDonough's (2014) findings that also revealed this effect. Neumann and McDonough (2014) explored opinion, descriptive, and comparison-contrast paragraphs written by EAP students following structured prewriting focusing on both content and organization. They found that students who preferred to work collaboratively and individually during prewriting generated better-quality texts, assessed through analytical ratings, after taking part in CP compared to the individual condition. Hence, this suggests that CP may be useful even for learners who prefer to perform tasks individually. Similar to McDonough et al.'s (2019) findings, a recent study by McDonough and De Vleeschauwer (2019) also detected the benefits of CP for accuracy and the lack of effect on complexity. Research compared the effects of CP and IP on the development of 60 Thai learners of English in a university over one semester. In this research, prewriting involved a review of background information, brainstorming, and outlining prior to three tasks: a data commentary, an opinion, and a process writing task. It was found that CP led to significantly improved accuracy over time compared to IP; however, no condition effects were observed on subordination or coordination complexity. Unlike McDonough et al.'s (2019) findings, however, higher analytical rating scores were detected in IP.

Other research in this area has yielded different findings. In another study by McDonough et al. (2018), they divided 128 Thai EFL learners into three groups, CP, collaborative writing, and no collaboration, and found that participation in an unstructured CP task did not lead to significant changes in the quality of problem-solution paragraphs produced afterwards. Scores for accuracy, subordination complexity, and each component of analytical ratings (content, organization, and language) of the prewriting and no-collaboration groups were comparable. Likewise, early work by Shi (1998) found no effects on ratings. Shi examined prewriting for English opinion essay-writing among 47 pre-university students incorporating three prewriting conditions: peer talk, teacher-led discussion, and no discussion. The differences across conditions in holistic rating scores were non-significant. However, the researcher observed that students tended to write longer texts under the no-discussion condition, possibly due to the increased amount of writing time available, while essays written after peer discussions tended to contain a larger variety of verbs. In contrast to this observation, the findings of Kang and Lee (2019)

revealed increased fluency in CP. The study explored the differences between individual and collaborative unstructured prewriting in 40 eighth-grade Korean learners of English, adopting picture narration writing tasks. Significant differences across the two prewriting conditions were detected for neither accuracy nor syntactic complexity, but increased fluency was evident under the collaborative condition. Interestingly, the influence of prewriting condition on lexical complexity was affected by task complexity; in complex writing tasks, the results were comparable across conditions, while language was lexically more complex under the individual planning condition in simple tasks. Other research that demonstrates complex relationships between the prewriting condition and other factors is that of Shin (2008). Shin investigated the influence of the prewriting condition and the interaction effects of this condition and proficiency on analytical ratings, assessed based on language, grammar, content, organization, and mechanics. Data obtained from argumentative and expository English essays written by Korean university students revealed that the collaborative group obtained significantly higher scores for all rating components than the individual group for expository writing; however, no differences were observed in argumentative writing and there were no interaction effects of the prewriting condition and proficiency.

In the realm of research exploring differences between various types of CP, Abrams and Byrd (2017) explored meaning- and grammar-focused collaborative prewriting and Pospelova (2021) investigated free discussion and structured prewriting. Abrams and Byrd (2017) compared texts written by participants performing grammar-focused (grammar review and editing) and meaning-focused (brainstorming of ideas and lexical items) activities and found that the meaning-focused group outperformed the other group on grammatical accuracy and lexical richness and obtained higher text-quality rating scores. In Pospelova's (2021) study, the researcher found that a structured prewriting task encouraging students to understand the writing task, brainstorm, evaluate/ organize ideas, and think about useful grammar and vocabulary led to more significant improvement than free prewriting discussions.

While most research has focused on performance following offline prewriting, a little has examined online prewriting in text chat. Liao (2018) compared the quality of Chinese texts produced by American students following unstructured face-to-face and text-chat CP. She found that while the texts generated after face-to-face discussions were significantly longer, those produced after text chat contained significantly more accurate characters and diverse vocabulary. In addition, there was evidence that prewriting in both modes may enhance writing fluency. Exploring the same topic, a small-scale study by Kessler et al. (2020) found that the Chinese opinion essays written by American learners post-offline discussion were significantly syntactically richer and lexically more complex than those generated after text chats. Another study by Amiryousefi (2017) compared three prewriting conditions: one individual planning condition and two text-chat collaborative planning conditions, teacher-monitored and student-led. Data from the two letter writing tasks performed by 76 intermediate learners of English revealed that different conditions may provide different benefits as regards writing fluency, accuracy, and complexity. Moreover, while the teacher-monitored condition facilitated the transfer of accuracy to the main writing task, individual planning resulted in the transfer of fluency.

In summary, previous L2 CP research has yielded mixed findings as regards writing performance. A small number of studies in this area have shown that not all collaborative tasks offer the same benefits, and the choices of prewriting tasks are important in determining writing outcomes. In addition, online studies comparing offline and text-chat prewriting have also yielded different findings concerning the benefits of each mode for performance.

Research questions

Owing to the limited amount of research exploring L2 collaborative prewriting, the present study aims to explore the potential of such planning for L2 learning guided by the following questions:

1. What are the effects of online prewriting conditions (individual vs collaborative voice-chat) on writing performance as assessed by text length, accuracy, and complexity?
2. What are the effects of online prewriting conditions (individual vs collaborative voice-chat) on writing performance as assessed by analytical ratings?

As previous findings pertaining to the effects of L2 collaborative prewriting on writing performance have been mixed, the research questions seek to understand its impact. Moreover, while, to the best of my knowledge, previous online L2 collaborative prewriting studies have focused solely on prewriting in text chat, this research investigates prewriting discussions in a new mode: voice chat. Because the two modes are inherently different, the present study might yield new knowledge about L2 online prewriting. Indeed, previous research on L2 collaborative writing has detected the benefits of voice chat (Cho, 2017; Elola & Oskoz, 2010; Oskoz & Elola, 2011, 2014). Incorporating it as a means for collaboration during writing, these studies found that learners perceived voice chat as an ideal way for the rapid exchange of ideas about content and organization (Elola & Oskoz, 2010; Oskoz & Elola, 2011), and they preferred voice chat over text chat due to its increased interactivity (Cho, 2017).

METHODOLOGY

Design

Data were gathered from 126 undergraduate Thai students in four online classes. Two classes were a control group (*n* = 66) performing the prewriting task individually, while the other two were an experimental group (*n* = 60) in the CP condition. All classes were conducted on Zoom (Zoom Video Communications, 2021), a video-conferencing platform. The texts students wrote during the main task were examined to answer the research questions. In compliance with the ethical standards of research involving human participants, this research project was approved by the ethics review committee of the university.

Participants

Participants were 126 Thai university students (59 females and 67 males) in the fields of

engineering and health sciences. They were enrolled on an online university English course targeting all four language skills that every first-year student was required to take. To ensure comparability, data were gathered from two classes of engineering students and two classes of health sciences students; each of these respective classes was randomly assigned to the experimental group. Participants had learned paragraph writing prior to data collection. Their L1 was Thai and their ages ranged from 18–22 years (*M* = 18.86, *SD* = .73). They had learned English for an average of 13.25 years (*SD* = 2.86, range = 6–19). Overall, they were intermediate writers based on teacher evaluation and their average score on a writing assignment was 10.27 (*SD* = 1.47, range = 7–12) out of 13. On data collection day, 83 (65.87%), 36 (28.57%), and 7 (5.56%) participants used a computer, a tablet, or a smartphone, respectively, when attending the online class. All reported using their device for at least half an hour per week with as many as 124 students (98.41%) reporting using it for more than three hours per week. No one reported not being familiar with their device.

Independent samples t-tests showed comparability between the control and experimental groups in the number of years spent learning English, $t(124) = 0.51$, $p = .614$, BCa 95% CI [-0.73, 1.35], and writing proficiency in terms of writing scores, $t(124) = 1.49$, $p = .139$, BCa 95% CI [-0.11, 0.64]. A Mann-Whitney U test also revealed comparability across groups in the amount of time spent per week on participants' selected devices, $U = 2094.00$, $p = .366$, $z = 0.90$, $r = .08$.

Main and prewriting tasks

The main task was problem-solution paragraph writing. The writing prompt, "Many people in Thailand do not recycle. Suggest two solutions to make them recycle more", was on a topic relevant to the unit covered in the course at that time. Students were required to write, individually, a paragraph of approximately 100–150 words including a topic sentence, supporting details, a concluding sentence, and transitional devices. All these components had been covered previously in the lesson.

The prewriting task was adapted from the structured prewriting tasks of McDonough et al. (2019) and Neumann and McDonough (2014, 2015). Two handout versions were created, one for each condition: IP and CP (see Appendix A). A main-task writing prompt was given in both versions, and each included two sections requiring learners to think about content and organization before writing. The first section concerned ideas generation. In this section, participants in the collaborative condition were asked to answer three questions relating to the writing prompt individually, then discuss their answers with a partner and provide feedback to him/ her. The second section concerned ideas selection and organization. In this section, participants in the collaborative condition were asked to consider the feedback received in the first part, individually select ideas to be incorporated in their paragraphs, write an outline, then discuss the outline with their partner. Again, they were instructed to provide feedback on their partner's ideas. The handout for the individual condition also contained the same two sections. The difference, however, was that instead of discussing ideas in pairs and providing feedback, participants in the individual condition were required to critically evaluate their ideas themselves. In essence, this research incorporates a meaning-focused structured prewriting task. As previous findings have suggested, compared to grammar-focused CP, meaning-focused

CP may yield more benefits in several aspects, i.e. grammatical accuracy, lexical richness, and overall text quality (Abrams & Byrd, 2017). Moreover, structured prewriting tasks potentially lead to increased gains in writing compared to unstructured ones (Pospelova, 2021).

Data collection procedure

Data collection took place in Week 10 of the semester during the normal three-hour class time. All classes were conducted on Zoom, the video conferencing software used by students since the start of the semester. The tasks were performed as part of the writing activities for the unit taken by students at that time. At the beginning, participants were told they would have an opportunity to complete a prewriting activity, either individually or collaboratively, before writing a problem-solution paragraph. A link to the prewriting handout, created on Google Forms, was then sent to them to complete and submit. At this stage, the teacher explained what students were required to do in the prewriting task and reminded them about what they had learned previously about paragraph and outline writing. Participants were instructed to complete the task by themselves without using dictionaries or online resources. Learners in the experimental group were sent to breakout rooms to complete the handout task in pairs, while the control group remained in the main Zoom conference room to complete their task. Breakout rooms are small conference rooms used to hold small group meetings. During online classes, participants in the main Zoom room can be sent to these rooms for small group discussions. In this research, two participants in the experimental condition were assigned to each room through Zoom's automatic breakout room generation function and, hence, group members were not predetermined. During discussions, 25 breakout rooms were also recorded through Zoom's recording function based on students' consent and the recording capability of their devices for the purposes of possible future research. Participants were given 40 minutes to complete the prewriting task and all were required to submit their handout individually. Before submitting, they were all instructed to save the outline on their device so that they could use it in the upcoming main task.

Upon completing the prewriting, the experimental group returned to the main room and both groups were sent a link to a writing platform, Testmoz (Aptibyte LLC, 2021). Students were asked to individually write and submit their paragraphs in 50 minutes. They were not allowed to use dictionaries or other resources during writing. This writing task was then followed by questionnaire completion; each participant was provided with a Google Forms link to complete a background questionnaire probing participants' background information and familiarity with the device they used and, for the purposes of a future study, an exit questionnaire probing their task performance experience and task perception.

Data analysis

Paragraphs composed in the individual-writing main task were examined through an analytic rubric and in three performance aspects: text length, accuracy, and complexity. Regarding the rubric (see Appendix B), two English language teachers trained by the researcher and experienced in using similar analytic rubrics rated the paragraphs independently based on content, organization, and language, each aspect on a scale of 0–5 points. Two-way mixed average-

measures intraclass correlation analyses run based on the absolute agreement model found excellent inter-rater reliability in all aspects: content, ICC = .94, 95% CI [0.91, 0.96], organization, ICC = .91, 95% CI [0.88, 0.94], and language, ICC = .94, 95% CI [0.92, 0.96]. The means of the content, organization, language, and total scores given by the two raters were subsequently used for statistical analyses.

Text length was measured by the number of words written, and accuracy by lexical, grammatical, and total errors/ word. All errors in the texts were identified and categorized as lexical or grammatical by a native speaker of English who was an experienced language teacher. Grammatical errors were those concerning grammatical mistakes, for instance, sentence structure, tense, voice, subject-verb agreement, singularity/ plurality, and word form. Lexical errors were any problems with word choice. Problems pertaining to spelling, capitalization, and punctuation were not coded as errors. Approximately 20% of the texts (n = 26), equally and randomly selected from the two groups, were then re-examined by the researcher to ensure coding reliability. Cohen's kappa coefficients show strong inter-coder agreement for both error identification (0.90) and categorization (0.83).

Finally, complexity was gauged in terms of both lexical and syntactic complexity. To ensure the validity of analyses, errors relating to spelling and punctuation were corrected prior to submitting the texts to automated complexity analysis tools. Different measures were adopted to investigate various dimensions of lexical complexity conceptualized in previous work including lexical density, sophistication, diversity, and disparity (Jarvis, 2013; Michel, 2017). Density was gauged by the ratio of content to total words obtained from VocabProfiler (Cobb, 2021), a Web-based vocabulary analysis tool. Sophistication was measured by the log frequency of content words, an index of lexical rarity (Jarvis, 2013) calculated utilising the CELEX lexical database (Baayen et al., 1995). Lexical diversity was assessed through the measure of textual lexical diversity (MTLD). This measure shows the average length of sequential word strings maintaining a given type/ token ratio value (McCarthy & Jarvis, 2010). Higher MTLD values represent larger proportions of unique words and, thus, increased diversity. Disparity, which pertains to "the degree of differentiation between lexical types in a text" (Jarvis, 2013, p. 25), was measured by Latent Semantic Analysis (LSA) following Jarvis' (2013) recommendation. This analysis captures similarity in semantic meaning by examining semantic overlap across sentences in a text. MTLD, LSA, and log frequency values were computed using Coh-Metrix 3.0 (Graesser et al., 2004, 2011; McNamara et al., 2014), an online text analysis tool.

Regarding syntactic complexity, following Norris and Ortega's (2009) suggestion, this research examined complexity at multiple levels: global, clausal, and phrasal. Globally, the texts were assessed based on syntactic similarity index and words/ T-unit. Similarity index indicates structural similarity by comparing the syntactic structures of all possible pairs of sentences. Texts with a high similarity index value are those with low syntactic complexity (Crossley et al., 2008). The second global measure, words/ T-unit, is a common SLA measure found to capture differences in L2 development and proficiency levels (Bulté & Housen, 2014; Ortega, 2003; Stockwell & Harrington, 2003). It indicates the average length of T-units, units defined as a main clause and all subordinate clauses embedded in or attached to it (Hunt, 1965). At the clausal level, the texts were examined for both coordination and subordination. Coordination

and subordination complexity were assessed by coordinate phrases/ clause and dependent clauses/ clause, respectively. Finally, phrasal complexity was assessed through words/ clause and complex nominals/ clause. Words/ clause is the phrasal complexity measure recommended by Norris and Ortega (2009). However, since clause length can be increased through insertion of adjuncts and other clause-level expansions and is perhaps not a purely phrasal complexity measure (Bulté & Housen, 2012), complex nominals/ clause was adopted as a complementary measure. According to Cooper (1976), complex nominals include 1) gerunds or infinitives in the subject position, 2) nominal clauses, and 3) nouns with any of the following modifiers: appositive, participle, possessive, prepositional phrase, adjective, and adjective clause. All syntactic complexity indices except syntactic similarity index were obtained from Synlex (Lu, 2010), a Web-based text analysis tool. Syntactic similarity index was calculated utilising Coh-Metrix 3.0.

Statistical analysis

To address both research questions, scores for each component, text length, accuracy, complexity, and analytical rating, were compared across groups through independent-sample t-tests utilising SPSS Version 22. Prior to these comparisons, the data were examined for outliers and normality and the correlations between dependent measures were tested. Most combinations of measures were found to be independent of each other with the highest r observed being .78. The only exception was the relationship between two accuracy measures, grammatical and total errors/ word, which were significantly and very strongly correlated ($p < .001$, $r = .92$)[1]. Following Plonsky and Oswald's (2014) standards for interpreting effect sizes in L2 studies, Cohen's d, with absolute values of 0.40, 0.70, and 1.00, was interpreted as indicating a weak, moderate, and strong effect, respectively.

RESULTS

Table 1 provides descriptive statistics for all measures and Table 2 shows t-test results for the comparisons across groups concerning these measures. Concerning the first research question, the t-test revealed that a significantly higher number of words was produced in the experimental condition and the effect size was large. In terms of accuracy and complexity, all t-test results were non-significant except for the measures of lexical density and phrasal complexity – content words/ total words and complex nominals/ clause. There was a significantly higher proportion of content words and more complex nominals/ clause in the experimental condition, indicating increased lexical and phrasal complexity. The effect size was small for the analysis concerning content-word proportion but in the medium range for complex nominals/ clause.

Regarding the second research question, all types of rating scores were significantly different across groups with higher scores observed in the experimental group. The effect sizes for these comparisons were in the small–medium range.

---

[1] In spite of this outcome of Pearson's correlation analysis, both grammatical and total errors/ word were included in further analyses because the number of total errors/ word shows an overall view of accuracy.

Table 1

Descriptive statistics for all measures

	Control (*n* = 66)		Experimental (*n* = 60)	
	M	*SD*	*M*	*SD*
Length *(words)*	137.61	20.81	157.67	46.17
Accuracy *(errors/ word)*				
grammar	0.07	0.04	0.06	0.03
lexis	0.014	0.01	0.012	0.01
total	0.08	0.04	0.07	0.03
Lexical complexity				
ratio of content to total words	0.53	0.05	0.55	0.05
log frequency of content words	2.37	0.12	2.35	0.13
MTLD	66.61	24.21	61.59	16.87
LSA index	0.32	0.11	0.35	0.11
Syntactic complexity				
syntactic similarity index	0.11	0.03	0.12	0.03
words/ T-unit	14.89	3.13	15.57	4.10
coordinate phrases/ clause	0.20	0.13	0.23	0.15
dependent clauses/ clause	0.33	0.12	0.32	0.14
words/ clause	9.81	1.55	10.43	2.38
complex nominals/ clause	1.12	0.31	1.31	0.39
Rating scores				
content	3.50	0.78	3.93	0.79
organization	3.19	0.88	3.58	0.92
language	3.36	0.84	3.70	0.83
total	10.05	2.08	11.21	2.04

Table 2

Effects of group

					BCa 95% CI	
	p	*d*	*t*	*df*	Lower	Upper
Length *(words)*	.003	0.96	-3.09	80.32	-34.29	-7.14
Accuracy *(errors/ word)*						
grammar	.097	-0.28	1.67	124	-0.002	0.02
lexis	.299	-0.19	1.04	124	-0.003	0.01
total	.067	-0.31	1.85	124	-0.001	0.03
Lexical complexity						
ratio of content to total words	.043	0.35	-2.04	124	-0.03	-0.001
log frequency of content words	.355	-0.17	0.93	124	-0.02	0.07
MTLD	.176	-0.21	1.36	116.39	-2.51	13.53
LSA index	.185	0.24	-1.33	124	-0.07	0.01
Syntactic complexity						
syntactic similarity index	.697	0.07	-0.39	124	-0.01	0.01
words/ T-unit	.295	0.22	-1.05	124	-2.03	0.56
coordinate phrases/ clause	.134	0.29	-1.51	124	-0.09	0.01
dependent clauses/ clause	.473	-0.14	0.72	124	-0.03	0.06
words/ clause	.087	0.40	-1.73	99.59	-1.33	0.09
complex nominals/ clause	.002	0.63	-3.10	124	-0.32	-0.07
Rating scores						
content	.003	0.54	-3.03	124	-0.72	-0.14
organization	.015	0.45	-2.47	124	-0.71	-0.04
language	.023	0.41	-2.30	124	-0.65	-0.04
total	.002	0.56	-3.16	124	-1.86	-0.42

DISCUSSION

This study compares the effects of L2 individual and collaborative prewriting on writing performance as assessed by text length, accuracy, complexity, and analytical ratings. The results reveal that the areas most influenced by the prewriting condition were analytical ratings and text length. Regarding ratings, the collaborative condition significantly outperformed the individual condition in all rating types: content, organization, language, and total rating score. To a certain extent, the findings corroborate the outcome of Elabdali's (2021) meta-analysis revealing that the texts learners wrote individually after having experienced collaborative writing received higher ratings compared to those written after experiencing individual writing. These findings are also in line with those of previous L2 research showing CP led to significantly higher content, organization, and language ratings than the individual condition (McDonough et al., 2019; Neumann & McDonough, 2014). Notably, as in the present research, these previous works suggesting the potential of CP over IP (McDonough et al., 2019; Neumann & McDonough, 2014) all adopted collaborative and individual structured prewriting activities focusing on content and organization. Hence, it is possible that pre-task planning of content and organization is more beneficial to learners who complete it collaboratively than individually. When learners have opportunities for interaction, they may receive input and feedback from their peers which can help them generate better texts in terms of content, organization, and language. When examining the recordings of participants' prewriting collaboration, it was found that participants did not discuss language or organization very much during CP but focused primarily on content. Hence, in terms of improved language and organization, it could be that, after generating ideas with a peer during CP, learners had more cognitive resources available to focus on these two aspects when writing individually, leading to improvements in these areas compared to the individual condition.

However, while this study found positive effects for CP, McDonough et al. (2018) who adopted an unstructured prewriting task observed a lack of differences in ratings between CP and non-collaborative groups. Such a discrepancy might be due to the difference in the prewriting task adopted; when prewriting is unstructured and learners can choose to plan freely, the activity may not be as effective in prompting desirable writing outcomes.

With regard to length, this study detected significantly more words written under the experimental compared to the control condition, with a large effect size. In other words, completing a prewriting task focusing on content and organization individually was not as effective in enhancing fluency as completing the same task collaboratively. This indicates the potential of collaboration during prewriting and corroborates the trend found in most previous research that has revealed, for instance, that CP led to increased number of words per minute compared to IP (Kang & Lee, 2019); prewriting collaboration both face-to-face and via text chat led to the production of longer texts over time (Liao, 2018); and student-led CP in text chat resulted in a higher average number of words compared to IP (Amiryousefi, 2017). Based on the observation of the recordings of participants' prewriting discussions and most previous studies that have detected learners' focus on content during L2 CP (e.g. Liao, 2018; McDonough et al., 2018, 2019; Neumann & McDonough, 2015), it may be that prewriting discussions facilitate ideas exchange and generation; this may, therefore, enable learners to spend less time on

conceptualization during the main individual writing task and more on writing, leading to the production of longer texts.

Turning to complexity, the analyses reveal that syntactic complexity at the global and clausal levels was not significantly different across conditions. This is in line with the findings of most previous studies suggesting a lack of significant differences between IP and CP in global (e.g. T-unit length), subordination (e.g. dependent clauses/ clause or main clause), and coordination (e.g. coordinated phrases/ clause, coordinated phrases and clauses/ word) complexity (Kang & Lee, 2019; McDonough & De Vleeschauwer, 2019; McDonough et al., 2018, 2019). As for the two measures of phrasal complexity, the two groups did not differ in clause length, corroborating the finding of Kang and Lee (2019); however, the experimental group generated significantly more complex nominals/ clause than the control group, with a medium effect size. The difference detected for this phrasal complexity measure indicates that the collaborative group potentially benefited more from prewriting than the individual group. Possibly, learners in the collaborative condition obtained more detailed information for their writing through prewriting discussions, which may have allowed them to describe and modify nouns more extensively through complex linguistic constructions. As regards lexical complexity, this study found lexical sophistication, diversity, and disparity to be comparable across conditions. Lexical density assessed by the ratio of content to total words was the only lexical measure demonstrating a significant effect of condition, with the experimental group producing a higher proportion of content to total words. While there has been a scarcity of research comparing lexical complexity in L2 IP and CP, Shi (1998), who explored three prewriting conditions, peer talk, teacher-led discussion, and no discussion, found a wider variety of verbs in essays written following peer talk. To some extent, this supports the significant finding of this study, indicating that lexical complexity might be promoted by prewriting collaboration. In the context of this study, the experimental group potentially generated a higher proportion of content words due to the increased number of ideas they aimed to convey. After discussing content with their peers during prewriting, they may have generated several ideas resulting in lexically denser texts.

Regarding accuracy, the analyses show non-significant outcomes for lexical, grammatical, and total error rates, supporting the findings of previous L2 studies (Kang & Lee, 2019; McDonough et al., 2018) that revealed no differences between CP and the non-collaborative condition for accuracy (e.g. errors/ T-unit or word). In their study, Kang and Lee (2019) hypothesized that this non-significant finding could be due to the limitations of the unguided prewriting task adopted to foster accuracy. In the present research, although learners were guided to plan their content and organize their ideas, this prewriting task focused more on ideas conceptualization and arrangement than language features, which potentially explains the non-significant findings observed pertaining to accuracy. Another possible explanation for the findings of this study is that the performance aspects influenced by CP may depend on learners' prioritization. Some previous L2 studies comparing IP and CP have found that CP positively influenced accuracy but not syntactic complexity (McDonough & De Vleeschauwer, 2019; McDonough et al., 2019), while some have demonstrated its positive effect on fluency, but not accuracy or syntactic complexity (Kang & Lee, 2019). In the present research, the collaborative group outperformed the individual group on text length and some lexical and syntactic complexity measures, but

not accuracy. It could be that CP does make the main individual writing task less taxing by reducing the cognitive load during the task, thereby enabling learners to attend more to certain language aspects, but what these aspects are might depend on learners' prioritization. Based on the trade-off hypothesis (Skehan, 1998), learners may not be able to attend to all aspects at the same time due to limited attention span.

Pedagogical implications, limitations, and future research

While some previous research has found benefits of offline L2 CP, the findings of this study suggest that adopting an L2 collaborative voice-chat prewriting task focusing on content and organization could also be useful; it may enhance performance in several aspects, rating scores (for content, organization, and language), complexity, and especially text length, when compared to IP. When teachers do not have sufficient time to let their students collaboratively compose a whole text together in online class, this prewriting might be an effective alternative task to implement. Second, the lack of effect on accuracy detected indicates the need for teachers to draw learners' attention to language during prewriting if the aim is to improve accuracy. Because participants preoccupied with content during discussions, if their attention had been directed towards language during CP, for example by requiring them to discuss lexical items and language structures that could be used in their texts, accuracy might be enhanced.

This research, however, also has some limitations. The first limitation is that it explores L2 prewriting among Thai university students and the findings may not be generalizable to other contexts. Second, it only adopted one type of prewriting task. Other prewriting activities could have yielded different outcomes. It would be useful for future research to explore performance by employing other types of prewriting or comparing various prewriting tasks. Third, the study investigated only one type of main task: problem-solution paragraph writing. The results could have been different if other task types had been adopted. Indeed, Kang and Lee (2019) found that task complexity levels potentially influence the effects of prewriting on lexical performance. Future studies could utilise a type of task different from that in this study, e.g. opinion or compare-contrast paragraph/ essay writing, or compare the results obtained from different task types or tasks with different degrees of complexity.

Owing to the limited amount of L2 CP research, more research is needed to understand its benefits for L2 learning and how best to implement this type of task. Within this field, researchers could focus on exploring L2 CP in online platforms as few studies have shed light on this, and this knowledge would allow educators to keep up with the ever-changing world where learning is increasingly migrating online. Moreover, more research comparing L2 CP and collaborative writing could be useful for educators. Other possible areas for L2 CP research concern cognitive processes, development over time, and perception. First, this research shows some positive effects of CP, potentially because it frees up resources in learners' working memory during the main task; exploring learners' cognitive processes during the main writing task, e.g. through stimulated recall, could further reveal whether the positive influences of this prewriting actually derive from the reduced cognitive load during some writing processes. To the best of my knowledge, no L2 prewriting research has examined cognitive processes. Regarding writing development over time and perception, studies investigating these could offer a better

understanding of the benefits of L2 CP over time and the connection between learners' perceptions and their prewriting discussions/ subsequent performance, which is currently lacking. Finally, researchers interested in L2 CP should consider exploring the impacts of different task conditions, e.g. group size or mode, on performance and prewriting discussions; through these investigations, educators will be able to make more informed decisions pertaining to prewriting task conditions.

CONCLUSION

This research was motivated by the limited number of studies exploring L2 CP and in an attempt to extend our understanding of online prewriting beyond text chat. It explores L2 prewriting in a novel mode, voice chat, by comparing the effects of CP and IP on writing performance. The findings show that learners might benefit more from completing prewriting tasks collaboratively rather than individually as the collaborative group significantly outperformed the individual group in various measures: text length, syntactic and lexical complexity, and ratings for language, organization, and content. These findings indicate the advantages of online collaboration during the prewriting stage. Learners can potentially conceptualize their ideas in greater detail when they have an opportunity to exchange ideas and receive peer feedback during voice-chat prewriting, thereby reducing the need for conceptualization and the cognitive load during writing, which in turn benefits text production.

THE AUTHOR

Vararin Charoenchaikorn, Ph.D., is an instructor at Chulalongkorn University Language Institute. Her research interests include second language acquisition, task-based language teaching, computer-assisted language learning, and second language writing.
cvararin@hotmail.com

REFERENCES

Abrams, Z. (2019). Collaborative writing and text quality in Google Docs. *Language Learning & Technology, 23*(2), 22–42. https://doi.org/10125/44681

Abrams, Z., & Byrd, D. R. (2017). The effects of meaning-focused pre-tasks on beginning-level L2 writing in German: An exploratory study. *Language Teaching Research, 21*(4), 434–453. https://doi.org/10.1177/1362168815627383

Amiryousefi, M. (2017). The differential effects of collaborative vs. individual prewriting planning on computer-mediated L2 writing: Transferability of task-based linguistic skills in focus. *Computer Assisted Language Learning, 30*(8), 766–786. https://doi.org/10.1080/09588221.2017.1360361

Aptibyte LLC. (2021). *Testmoz*. https://testmoz.com/

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. Linguistic Data Consortium. https://doi.org/10.35111/gs6s-gm48

Beauvois, M. H. (1997). Computer-mediated communication (CMC): Technology for improving speaking and writing. In M. D. Bush & R. M. Terry (Eds.), *Technology-enhanced language learning* (pp. 165–184). National Textbook Company.

Bikowski, D., & Vithanage, R. (2016). Effects of web-based collaborative writing on individual L2 writing development. *Language Learning & Technology, 20*(1), 79–99. https://doi.org/10125/44447

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins. https://doi.org/10.1075/lllt.32.02bul

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Cho, H. (2017). Synchronous web-based collaborative writing: Factors mediating interaction among second-language writers. *Journal of Second Language Writing, 36*, 37–51. https://doi.org/10.1016/j.jslw.2017.05.013

Chun, D. M. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System, 22*(1), 17–31. https://doi.org/10.1016/0346-251X(94)90037-X

Cobb, T. (2021). *Web Vocabprofile*. https://www.lextutor.ca/vp/

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research, 69*, 176–183. https://doi.org/10.1080/00220671.1976.10884868

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly, 42*, 475–493. https://doi.org/10.1002/j.1545-7249.2008.tb00142.x

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*. https://doi.org/10.1016/j.jslw.2020.100788

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy of L2 oral production. *Applied Linguistics, 30*(4), 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*(1), 59–84. https://doi.org/10.1017/s0272263104261034

Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology, 14*(3), 51–71.

Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*, 40–58. https://doi.org/10.1016/j.jslw.2011.12.002

Flower, L., & Hayes, J. R. (1981). Plans that guide the composing process. In C. H. Frederiksen, M. F. Whiteman & J. F. Dominic (Eds.), *Writing: The nature of development, and teaching of written communication (Vol.2)* (pp. 365–387). Erbaum Associates.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Ghavamnia, M., Tavakoli, M., & Esteki, M. (2013). The effect of pre-task and online planning conditions on complexity, accuracy, and fluency on EFL learners' written production. *Porta Linguarum, 20*, 31–43. https://doi.org/10.30827/Digibug.29559

Graesser, A. C., McNamara, D. S., & Kulikowich, J. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*, 223-234. https://doi.org/10.3102/0013189X11413260

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers, 36*, 193–202. https://doi.org/10.3758/BF03195564

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 29–55). Lawrence Erlbaum Associates.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). National Council of Teachers of English. https://files.eric.ed.gov/fulltext/ED113735.pdf

Jarvis, S. (2013). Defining and measuring lexical diversity. In S. Jarvis & M. Daller (Eds.), *Vocabulary knowledge: Human ratings and automated measures* (pp. 13–44). John Benjamins. https://doi.org/10.1075/sibil.47.03ch1

Joaquin, A. D. L., Kim, S. H., & Shin, S.-Y. (2016). Examining prewriting strategies in L2 writing: Do they really work? *Asian EFL Journal, 18*(2), 156–189.

Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing, 21*, 264–282. https://doi.org/10.1016/j.jslw.2012.05.011

Kang, S., & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing, 45*, 61–72. https://doi.org/10.1016/j.jslw.2019.08.001

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(2), 355–365. https://doi.org/10.1037/0278-7393.14.2.355

Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *The American Journal of Psychology, 103*(3), 327–342. https://doi.org/10.2307/1423213

Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *The Modern Language Journal, 79*(4), 457–476. https://doi.org/10.2307/329999

Kessler, M., Polio, C., Xu, C., & Hao, X. (2020). The effects of oral discussion and text chat on L2 Chinese writing. *Foreign Language Annals, 53*(4), 666–685. https://doi.org/10.1111/flan.12491

Kroll, B. (2002). Considerations for teaching an ESL/EFL writing course. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 219–232). Heinle & Heinle.

Lan, Y. J., Sung, Y. T., Cheng, C. C., & Chang, K. E. (2015). Computer-supported cooperative prewriting for enhancing young EFL learners' writing performance. *Language Learning & Technology, 19*(2), 134–155.

Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 171–198). John Benjamins. https://doi.org/10.1075/lllt.32.08lev

Liao, J. (2018). The impact of face-to-face oral discussion and online text-chat on L2 Chinese writing. *Journal of Second Language Writing, 41*, 27–40. https://doi.org/10.1016/j.jslw.2018.06.005

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

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

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*, 381–392. https://doi.org/10.3758/BRM.42.2.381

McDonough, K., & De Vleeschauwer, J. (2019). Comparing the effect of collaborative and individual prewriting on EFL learners' writing development. *Journal of Second Language Writing, 44*, 123–130. https://doi.org/10.1016/j.jslw.2019.04.003

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

McDonough, K., De Vleeschauwer, J., & Crawford, W. (2019). Exploring the benefits of collaborative prewriting in a Thai EFL context. *Language Teaching Research, 23*(6), 685–701. https://doi.org/10.1177/1362168818773525

McDonough, K., & González, T. H. (2020). Collaboration before writing: Exploring how student talk contributes to English L2 written narratives. *REFLections, 27*(1), 62–80.

McDonough, K., & Neumann, H. (2014). Using prewriting tasks in L2 writing classes: Insights from three experiments. *TESL Canada Journal, 31*, 132–143. https://doi.org/10.18806/tesl.v31i0.1190

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press. https://doi.org/10.1017/CBO9780511894664

Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50–68). Taylor & Francis.

Neumann, H., & McDonough, K. (2014). Exploring the relationships among student preferences, prewriting tasks, and text quality in an EAP context. *Journal of English for Academic Purposes, 15*, 14–26. https://doi.org/10.1016/j.jeap.2014.05.002

Neumann, H., & McDonough, K. (2015). Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing. *Journal of Second Language Writing, 27*, 84–104. https://doi.org/10.1016/j.jslw.2014.09.009

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

Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System, 34,* 566–585. https://doi.org/10.1016/j.system.2006.08.003

Ong, J., & Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Quarterly, 47*(2), 375–398. https://doi.org/10.1002/tesq.55

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Oskoz, A., & Elola, I. (2011). Meeting at the wiki: The new arena for collaborative writing in foreign language courses. In M. J. W. Lee & C. McLoughlin (Eds.), *Web 2.0-based e-learning: Applying social informatics for tertiary teaching* (pp. 209–227). IGI Global. https://doi.org/10.4018/978-1-60566-294-7.ch011

Oskoz, A., & Elola, I. (2014). Promoting foreign language collaborative writing through the use of web 2.0 tools and tasks. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 115–148). https://doi.org/10.1075/tblt.6.05osk

Plonsky, L., & Oswald, F. L. (2014). How big is "Big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Pospelova, T. (2021). The collaborative discussion model: Developing writing skills through prewriting discussion. *Journal of Language and Education, 7*(1), 156–170. https://doi.org/10.17323/jle.2021.10748

Rostamian, M., Fazilatfar, A. M., & Jabbari, A. A. (2018). The effect of planning time on cognitive processes, monitoring behavior, and quality of L2 writing. *Language Teaching Research, 22*(4), 418–438. https://doi.org/10.1177/1362168817699239

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

Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *Journal of Second Language Writing, 7*(3), 319–345. https://doi.org/10.1016/S1060-3743(98)90020-0

Shin, Y. (2008). *The effects of planning on L2 writing: A study of Korean learners of English as a foreign language* (Publication No. 3315989) [Doctoral dissertation, The University of Iowa]. ProQuest Dissertations & Theses.

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

Stockwell, G., & Harrington, M. (2003). The incidental development of L2 proficiency in NS-NNS email interactions. *CALICO Journal, 20*(2), 337–359. http://www.jstor.org/stable/24149502

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., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Multilingual Matters.

Strobl, C. (2014). Affordances of Web 2.0 technologies for collaborative advanced writing in a foreign language. *CALICO Journal, 31*(1), 1–18. https://doi.org/10.11139/cj.31.1.1-18

Sullivan, N., & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *System, 29*(4), 491–501. https://doi.org/10.1016/S0346-251X(96)00044-9

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of Henry G. Widdowson* (pp. 125–144). Oxford University Press.

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

Wang, Y. C. (2015). Promoting collaborative writing through wikis: A new approach for advancing innovative and active learning in an ESP context. *Computer Assisted Language Learning, 28*(6), 499–512. https://doi.org/10.1080/09588221.2014.881386

Weigle, S. C. (2014). Considerations for teaching second language writing. In M. Celce-Murcia, D. Brinton & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 222–237). Heinle Cengage Learning.

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26*, 445–466. https://doi.org/10.1177/0265532209104670

Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. McGraw-Hill.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1–27. https://doi.org/10.1093/applin/24.1.1

Zoom Video Communications. (2021). *Zoom*. https://zoom.us/

Appendix A

Collaborative prewriting task handout

You will individually write a paragraph of approximately 100–150 words that includes topic, supporting, and concluding sentences and transitional devices on the following topic:

> Many people in Thailand do not recycle. Suggest two solutions to make them recycle more.

You must suggest two practical solutions to the problem and support your solutions with logical and relevant supporting details.
Before you start writing, you will discuss ideas with a partner by following the steps below.

Step 1: Generating ideas

1.1) Look at the 3 questions (a–c) below and individually type your answers in the space provided under each question. You do not have to write in complete sentences at this stage and please do not discuss your answers with your friend yet.

> a. Why do you think many people in Thailand do not recycle?
>
> b. What are the negative effects of the low recycling rate in Thailand?
>
> c. What are potential solutions that will make Thai people recycle more?

1.2) Once you have answered all the questions (a–c), discuss your answers with your partner. While discussing, you may share your screen so your partner can see your answers. You can add more ideas to your list while you are talking or delete any ideas that you do not like.

> * As you listen to your partner's ideas, tell your partner whether you agree or disagree with him/her and provide reasons for your opinion.

Step 2: Selecting and organizing ideas

2.1) Consider the feedback you received during 'Step 1' and individually select the ideas to be included in your paragraph, then write an outline of your paragraph in the space below. You do not have to write in complete sentences at this stage. Please do not discuss with your partner yet.

2.2) Once you have completed the outline, discuss it with your partner to receive feedback and make revisions to improve your outline. While discussing, you may share your screen so your partner can see your outline.

* As you listen to your partner's ideas, tell him/her your opinion. Provide feedback on his/her outline and make suggestions whenever possible.

Appendix B

Analytic rubric

Score	Criteria		
	Content	Organization	Language
5	- Clearly stated problem - Two clear solutions supported with reasonable, relevant, and sufficient details; very well-developed ideas	- Containing all 3 components[1] - Ideas linked appropriately with accurate use of a variety of linguistic devices; the paragraph read smoothly throughout	- Containing a good variety of sentence structures and lexical items - Containing no major errors[2] and no/ very few minor errors[3] that do not affect comprehensibility
4	- Clearly stated problem - Two clear solutions supported with reasonable, relevant, and sufficient details	- Containing all 3 components - Ideas linked together although with a few errors (e.g. a few transitional devices may be lacking, overused, or used inappropriately) OR - Lacking one component - Ideas linked appropriately with accurate use of a variety of linguistic devices; the paragraph read smoothly throughout	- Containing a good variety of sentence structures and lexical items, yet wordy or repetitive in a few places - Containing a few major and minor errors
3	- Clearly stated problem - One solution not clearly stated or lacking logical, relevant, and sufficient supporting details	- Containing all 3 components - Many ideas not linked together appropriately, making the text difficult to follow OR - Lacking one component - Ideas linked together although with a few errors (e.g. a few transitional devices may be lacking, overused, or used inappropriately)	- Containing an adequate variety of sentence structures and lexical items - Containing occasional major and minor errors

Score	Criteria		
	Content	Organization	Language
2	- Clearly stated problem - Two solutions not clearly stated or lacking logical, relevant, and sufficient supporting details OR - No clearly stated problem - One solution not clearly stated or lacking logical, relevant, and sufficient supporting details	- Containing all 3 components - Minimal organization and use of transitional devices OR - Lacking one component - Many ideas not linked together appropriately, making the text difficult to follow	- Containing a limited range of sentence structures and lexical items, showing poor language ability - Containing frequent major and minor errors
1	- No clearly stated problem - Two solutions not clearly stated or lacking logical, relevant, and sufficient supporting details OR - No clearly stated problem - Containing only one solution	- Lacking one component - Minimal organization and use of transitional devices OR - Lacking more than one component; incomplete paragraph	- Containing a very limited range of sentence structures and lexical items, showing very limited language ability - Containing extensive major and minor errors, showing the writer's lack of control over accuracy
0	- The paragraph is completely incomprehensible or not related to the prompt.		

Note.

[1] Main components: topic sentence, supporting sentences, and conclusion

[2] Major errors (seriously affecting comprehension), e.g. tense, voice, sentence fragment, run-on sentence, word form, and word choice

[3] Minor errors (not seriously affecting comprehension), e.g. article, spelling, capitalization, and punctuation