
Exploring the use of collaborative writing in an EFL classroom context

ANNITA STELL

The University of Queensland

ABSTRACT

To date, the use of collaborative writing remains limited in second language (L2) classrooms despite the promising results shown in past studies. From a sociocultural theory (SCT) perspective, co-authored tasks can push learners to use language as a tool to pool L2 resources together or scaffold their writing (Storch, 2013). Little research, however, has examined the effects paper-based collaborative writing tasks have on EFL learners' writing. Building on past research, this study aims to investigate the effects of collaborative writing on complexity, accuracy and fluency (CAF) in the writing of Taiwanese learners and on patterns in their collaborative interactions about their writing. Thirty-two Taiwanese university students completed two writing tasks: one individually and one in pairs. A mixed-methods approach was used. The individual and pair texts were quantitatively compared for levels of CAF, and pair talk was qualitatively analysed for language-related episodes (LREs) and patterns of interaction. The quantitative findings showed that pairs who worked in a collaborative manner were able to achieve better accuracy and fluency. The findings also showed how collaborative writing enabled writers to resolve meaning-related issues in their texts.

Address for correspondence: Annita Stell, The School of Languages and Cultures, Room 320, Third Floor, Gordon Greenwood Building (32), The University of Queensland, St Lucia, QLD 4072, Australia; Email: a.stell@uq.edu.au

University of Sydney Papers in TESOL, 13, 63-97.

©2018 ISSN: 1834-4712 (Online)

INTRODUCTION

The ability to write is a skill many students struggle to develop, especially in English-as-a-foreign-language (EFL) contexts such as Taiwan. When the target language is not readily available or commonly used outside predominantly exam-oriented classrooms, having to apply various writing conventions relating to genre, structure and linguistic knowledge appropriately can be overwhelming (Philp, Adams, & Iwashita, 2014). In such classrooms, quizzes and final exams are viewed as reliable tools to measure the students' level of progress (Shohamy, 2017). The significant number of tests involved in their schooling shifts students' priorities towards rote memorisation and test preparation strategies to achieve higher grades.

Like other Asian cultures influenced by Chinese Confucianism, academic study in Taiwan is highly valued and respected (Chou & Yuan, 2011). To gain (or maintain) high grades, *Buxibans* (or cram schools) that students attend after school, remain popular (or perhaps, crucial) to all stake-holders including parents, teachers and students. True to its name, cram schools tend to focus on drilling learning content through the use of test-taking strategies (Chung, 2013).

This strong test-taking orientation is also reflected in Taiwan's Ministry of Education (MOE) reform policies for English language learning, which aims to support the development the nation's English language proficiency. For example, in the 2008 Administration Guidelines, attention was placed on encouraging universities to have exit English benchmark requirements that reflect a B1 or A2 standard from the Common European Framework of Reference for Languages (CEFR) (Chu & Yeh, 2017; Pan & Newfields, 2012). Over the years, MOE has continued to promote studying abroad, which has increased the need for standardised proficiency tests. A few studies, however, have voiced their concerns about the negative washback these testing conditions can create if the focus remains on teaching and learning for tests (Chung, 2013; Pan, 2015; Wu, 2011).

Because of issues such as these, exploring learning approaches that increase L2 usage is a key area of interest. Currently, many

researchers are focussing on finding ways to shift writing from being a solitary act to a social act involving peer interaction that encourages learners to verbalise and exchange their ideas (Swain, 2013). This shift is also emphasised in Taiwanese MOE's objectives for 2019: train teachers to implement "student-centred learning practices" (Ministry of Education Republic of China, 2018, para 5).

In particular, with the rise of technology and interactive learning, collaborative writing has gained attention in L2 writing pedagogy (Elola & Oskoz, 2010). Collaborative writing tasks are defined as tasks in which two or more authors co-construct texts. Storch (2013) believes that there are two characteristics that distinguish these from cooperative tasks. While both involve two or more authors, a collaborative task requires meaningful interaction and shared-decision making (Storch, 2013), whereas a cooperative task gives authors the option to allocate different sections to different authors and compile individually written work into a single document.

In recent years, research on co-constructed texts has shifted from a traditional pen and paper approach to a multi-modal computer-assisted approach (Li & Storch, 2017). A lot of the existing studies examine how collaborative writing can be effective through various mediums (i.e., computers, tablets) for a range of written genres to increase L2 usage, autonomy and improve grammatical accuracy (Dobao, 2012). Existing research focusing on Taiwanese contexts tends to spotlight high-tech classrooms (e.g., Liao, Chang & Chan, 2018; Wang, 2015). Few studies explore traditional classroom settings.

Clearly, research exploring the use of computer-assisted learning is valuable, but it is still important to investigate classroom dynamics within traditional classrooms. This is because, from a Socio-cultural Theory (ST) perspective, language is mediated through the social context to help internalise psychological tools, such as language and signs (Lantolf & Poehner, 2011; van Compernelle & Williams, 2013). When learners spend long hours before, during and after-school cramming to excel in their written exams, time to engage in meaningful L2 conversations is limited. Consequently, the question of whether collaborative writing can be effective in exam-oriented

classroom settings is a valuable one to explore. In Taiwan specifically, only a small number of studies have examined how traditional paper-based collaborative writing can be implemented into EFL writing instruction. More research is needed to understand how the verbalisation of ideas can work as a mechanism that moves social interactions into the internal plane of psychological functioning (Swain, 2000). This study aims to provide a holistic examination of how collaborative writing can be implemented in Taiwan; it focuses specifically on understanding whether the communication that occurs between writers completing collaborative writing tasks helps to develop levels of complexity, accuracy and fluency (CAF).

THE EFFECTS OF COLLABORATIVE WRITING

In past studies, collaborative writing tasks were generally used as prompts to encourage pairs or groups to communicate purposely with each other without support from external resources (Cumming, 2012; Haneda, 2007). L2 production was not considered as the outcome of interactive or collaborative dialogue; instead, researchers focused on exploring how learners use language as a tool in various contexts to stretch their own language boundaries and create language learning opportunities through co-constructing or scaffolding that would gradually shift to self-regulation (Donato, 1994; Philp *et al.*, 2014).

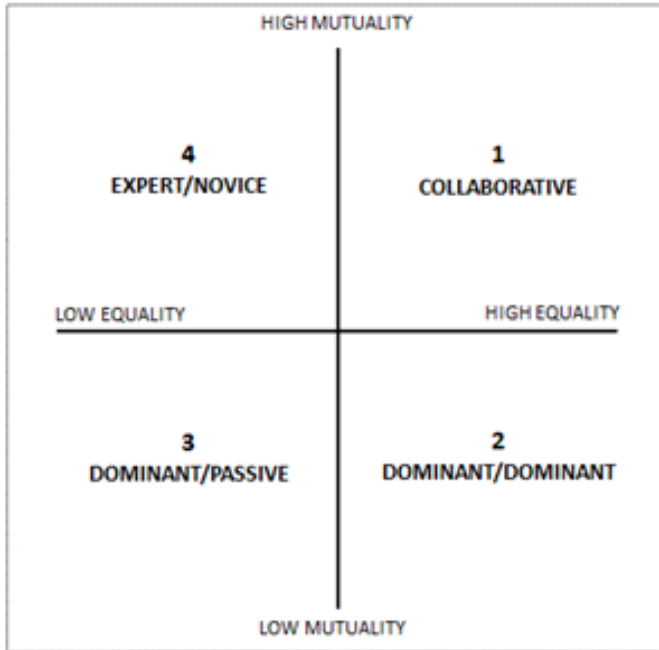
However, in recent years it has become common in SLA to analyse the impact collaborative dialogue has on actual writing development. For example, studies have used qualitative methods to analyse language-related episodes (LREs), which are episodes that focus on linguistic issues, such as spelling, grammar and syntax (Storch, 2008), along with quantitative methods to analyse written text quality. Swain (2006) emphasised that the process of languaging (or discussing language) can encourage learners to focus more on fine-tuning their language output and on experimenting with language. These discussions can also encourage learners to reach the zone of proximal learning development or ZPD, which is the distance between what individuals can do without assistance and what they can do with assistance; in other words, when learners interact with an expert or a more capable peer, they may have the opportunity to achieve beyond their individual capabilities (Vygotsky, 1987).

The impact of pair allocations has also been explored to understand the effect pairings have on task performance. Many researchers have compared the performance of different proficiency pairings. For example, Storch and Aldosari (2013) found that there were learning opportunities involved in similar proficiency pairings, even though many studies tend to emphasise the positive effects of expert-novice pairings. This may be because each individual has different strengths and weaknesses, so they can assist each other at different stages of the writing task; for instance, one may be more familiar with grammar rules, while the other is more confident with text structure. When learners with similar proficiency levels pool their linguistic knowledge together, they have the potential to achieve a higher performance level that is beyond their individual abilities (Dobao, 2012; Mirzaei & Eslami, 2015).

Storch (2002) identified four main patterns of dyadic interactions, which varied in terms of levels of equality (contribution) and mutuality (engagement). Equality refers to whether the contribution provided by each learner was equal, while mutuality refers to how well pairs stay on task and acknowledges each other's ideas. The four patterns were: collaborative, expert-novice, dominant-dominant, and dominant-passive (see Figure 1). Storch found that pairs mainly worked in collaborative, expert-novice or dominant-passive relationships. Other studies, such as Watanabe and Swain (2007), also confirmed that students tend to work in one of these three main patterns of interaction, but emphasised that collaborative interaction patterns had a stronger effect on overall task performance. When students worked collaboratively, they were able to contribute equally to the task and assist each other when needed.

FIGURE 1

Patterns of interaction (Storch, 2002)



A study by Tan *et al.* (2010) extended this body of research by comparing peer interactions through two mediums: face-to-face (FTF) and computer-mediated communication (CMC). The study highlighted how the same learners engaged differently, depending on the medium. For instance, the findings showed how learners engaged in additional cooperative patterns of interaction when they interacted through CMC. This pattern had similar characteristics to dominant-dominant patterns of interaction, as both learners displayed moderate to high equality levels and low mutuality levels. The difference between dominant-dominant and cooperative interactions, however, was shown through the way students divided the labour to complete the task. Cooperative pairs tended to focus more on

working on sections independently (with their own ideas) before compiling everything into a single document. In contrast, dominant-dominant pairs focused on following and expanding their own ideas while they both worked on the task together, which could at times lead to conflict. Working co-operatively, gave students a chance to contribute equally, even though they only completed their own sections. More research in this area would provide valuable information on how different patterns of interaction can be conducive to the development of writing skills.

There is a growing interest in examining the effects of collaborative writing on writing performance through analysing the levels of complexity, accuracy and fluency, along with overall writing quality (e.g., Watanabe & Swain, 2007; Wigglesworth & Storch, 2009). These studies examined how collaborative writing could enhance text quality through encouraging learners to discuss and resolve language-related issues without teacher assistance. Notably, many of these studies compared collaborative text quality with individual text quality (Dobao, 2012; Glendinning & Howard, 2003), and illustrated how collaborative writing activities can be used effectively as a means of learning to write (Weissberg, 2000; Sajedi, 2014). Weissburg (2000) explained that peer interaction could encourage learners to produce more varied syntactic forms when they are writing; it provides more opportunities for individuals to experiment with language. Pair and group work also encourage learners to solve linguistic problems to produce a more accurate text collaboratively (Wigglesworth & Storch, 2009).

Building on past research, the current study aims to explore the use of paper-based, collaborative writing in a Taiwanese context. The following research questions will be addressed:

1. How does writing performance differ between collaborative writing tasks and individual tasks?
2. How does the level of interaction and engagement within pairs affect L2 interaction patterns?

CONTEXT AND PARTICIPANTS

This study was a part of a larger study on the use of collaborative writing in EFL classrooms (author). Data was collected from thirty-two Taiwanese students who were enrolled in an exam-oriented IELTS preparation program that aimed to help students prepare for IELTS through practicing possible writing topics and memorising various sentence patterns. The students who signed up to participate in this study wanted to gain additional feedback on their general English writing ability, as well as more exposure to the target language, since this is not readily available outside of their language classrooms. In order to gain a general idea of the students' L2 knowledge in regards to grammar and spelling, an institutional version of the TOEFL test was used. TOEFL was used to avoid possible practice effects due to familiarity with the IELTS testing format.

RESEARCH PROCEDURES AND MATERIALS

Since students came from different IELTS preparation courses and language institutes, the data collection was conducted at a local language school in Kaohsiung City, Taiwan over two weeks.

Four picture ordering tasks involving Garfield cartoons were selected for this study: two individual and two collaborative tasks. Each one of the four tasks had the same instructions: rearrange the four pictures in the correct order and write a short paragraph about the pictures in the past tense (Appendix C). The four picture ordering tasks were based on familiar scenarios to ensure that each student could complete the task. Each participant was instructed to complete two of the four writing tasks - one individually (in 20 minutes) and one in pairs (in 40 minutes). To compare the individual and collaborative task quality, the order of tasks (individual/pair or pair/individual) and task prompts was counterbalanced. Table 1 shows the counterbalanced design.

Pilot test

The writing tasks were piloted by four students to determine the clarity of the instructions as well as the level of difficulty. The students found the instructions clear and the tasks manageable, but thought the

timeframe was too long. Hence, small changes were made: a word limit (100-150) was included, but the timeframe of 20 mins for the individual task and 40 mins for the collaborative task were kept the same, as some students may require more time than others.

TABLE 1
Counterbalanced design

	Session 2A		Session 2B	
	Group 1 (n=8)	Group 2 (n=8)	Group 3 (n=8)	Group 4 (n=8)
Task	1A	1B	2A	2B
	2B	2A	1B	1A

Session 1: Preparation

The purpose of the first data collection session was: (1) to gain a general indication of the students' linguistic knowledge, and (2) to familiarise the students with the kinds of writing tasks to be used in the study. Students were given a 30-minute, institutional version of the TOEFL test and two short practise writing tasks. Section 2: Structure and Written Expression, was used from the TOEFL test to determine their general proficiency level.

After the test, students were given two short activities similar to the actual collaborative and individual tasks to complete. Each student was given 10 minutes for the individual writing task and 15 minutes for the collaborative writing task. This was to ensure that the participants understood how to complete the tasks the following week.

Pair Allocations

Prior to the second session, the TOEFL tests were marked and converted according to Appendix A; the converted scores ranged from 27-59 with a mean score of 47.5. The participants' level of English ranged from CEFR A1-B2 (Appendix B), which means that the participants were either basic or independent users of English. Only one student achieved A1 with a score of 27, while most of the students (n=18) achieved B1; a similar number of students received

A2 (n=6) and B2 (n=5). Even though the TOEFL test did not assess all aspects of written communication, it gave a general indication that students had the knowledge needed for basic to independent usage of English. The students were allocated into either mixed proficiency or similar proficiency pairings. Appendix B shows the pairs using pseudonyms and their writing scores.

Session 2: Writing Tasks

Participants were given two writing tasks: one to be completed individually and the other in pairs. Different cartoons were given for the two writing tasks, according to the aforementioned counterbalanced design (see Table 1). It is important to note that this picture ordering tasks were considered unfamiliar to the students, as they normally focused on preparing for the academic IELTS writing tasks through rote memorisation and other test-taking strategies. To ensure students had enough time to complete the tasks, 20 minutes were given for the individual task, while 40 minutes was allocated for the pair task to ensure that students had enough time to discuss and resolve issues.

During the pair task, students were instructed to record their discussions with their mobile phones. Mobile phones were placed on the desks to reduce sound distortion and discourage participants from using them as a form of assistance. Participants were also encouraged to discuss in their preferred language: English (L2), Mandarin and/or Taiwanese to encourage them to exchange ideas and work together to solve any language issues without external assistance.

DATA ANALYSIS

The data set included a total of 48 texts (thirty-two individual texts and sixteen pair texts) and fifteen sets of transcribed pair talk. However, as will be explained later, one pair was excluded. Thus, the data set was from a total of 15 pairs (n=30) (see Appendix B).

Research question 1

The first research question was addressed by analysing and comparing the individual and collaborative texts' levels of complexity, accuracy and fluency (CAF). While CAF remains a

common measure for language performance, no single definition of CAF exists (Norris & Ortega, 2009). In this study, complexity refers to the learners' ability to use a variety of vocabulary and structures in the target language (Housen *et al.*, 2012) while accuracy is defined as the text or speech that is free from error (Kuiken & Housen, 2009). Fluency however, refers to the smooth and effortless use of language (Housen *et al.*, 2012). It focuses on flow and level of automaticity, which is the learners' ability to process language without attention or control (Kowal, 2014). The following measures of CAF were selected CAF.

Clauses and Dependent Clauses

A main clause contains a subject and verb to create a complete thought; it can stand alone to convey meaning. A dependent clause however, can be described as an incomplete thought adds additional meaning to a clause (Wolfe-Quintero, Inagaki, & Kim, 1998).

T-units and Error-free T-units

T-units contain a main (independent) clause and subordinate clauses (Wolfe-Quintero *et al.*, 1998). An error-free T-unit is simply a T-unit without any errors or mistakes.

Example: |The dog jumps over the log|. (1 T-unit)

|The dog jumps| and |the log breaks|. (2 T-units)

Table 2 summarises these CAF measures. These CAF measures were analysed using both descriptive and inferential statistics to determine whether there was a difference in the students' collaborative and individual writing performance.

The effect size of each CAF measure was also calculated through Pearson's correlation, r to quantify the difference between the individual and collaborative written texts. The following guidelines were used to interpret the effect sizes: 0.1 (small effect size), 0.3 (medium effect size) and 0.5 (large effect size).

TABLE 2
Summary of CAF measures

Linguistic Complexity	<ul style="list-style-type: none"> ○ Clauses per T-units (C/T) ○ Dependent clauses per T-units (DC/T)
Linguistic Accuracy	<ul style="list-style-type: none"> ○ Error-free T-units per T-unit (EFT/T) ○ Errors per T-unit (E/T)
Fluency	<ul style="list-style-type: none"> ○ Counting the total number of words written ○ (W) ○ Number of words per error-free T-unit ○ (W/EFT)

Multiple pair sample t-tests were used to compare the CAF measures for both individual and collaborative texts. A preliminary analysis was first completed to identify potential outliers and conduct a normality test. Pair 16 were identified as outliers as they wrote very little and spent less than 10 minutes on each task; because of this, they were omitted from the data set. The results of the Shapiro-Wilk tests determined that the distribution of pair value differences, $T_{diff.}$ ($p=0.68$), $W/EFT_{diff.}$ ($p=0.29$), $C/T_{diff.}$ ($p=0.87$), $DC/T_{diff.}$ ($p=0.91$), $EFT/T_{diff.}$ ($p=0.79$), and $E/T_{diff.}$ ($p=0.089$), were approximately normally distributed ($p>0.05$).

Research question 2

For the second research question, to determine the effect pair talk had on learning opportunities, the transcribed dialogues from the collaborative texts were qualitatively analysed for (1) Language Related Episodes (LREs) frequencies and (2) patterns of interaction at all stages of writing (i.e., planning, writing and editing).

LRE frequencies

Table 3 shows that the LREs were coded according to three types of language functions: Lexical-focused (word choices; definitions), Mechanics-focused (punctuation; spelling) and Form-focused (grammar). These LREs were marked as resolved (✓), incorrectly resolved (✗) or unresolved (?). The instances of different language choices (English, Chinese or Taiwanese) were also counted to determine how choice of language (L1 or L2) affected how

participants resolved certain linguistic issues. Examples of correctly resolved, incorrectly resolved and unresolved LREs are then provided.

TABLE 3

Summary of LRE types

Type of LRE	Code	Focus
Lexis-focused	L-LRE	<ul style="list-style-type: none"> ○ Definitions ○ Word choices
Form-focused	F-LRE	<ul style="list-style-type: none"> ○ Grammar related
Mechanics-focused	M-LRE	<ul style="list-style-type: none"> ○ Spelling ○ Punctuation

Excerpt 1. Pair 13, Amy and Jenny, are focusing on the correct tense for 'fall'. This was coded as correctly resolved.

- 142 J: Soon fell asleep.
 143 A: Yah. Soon.
 144 J: Oh...Felt or fell?
 145 A: Fall.
 146 J: Oh Fall...
 147 A: Fall asleep, the past tense is fell asleep. And had a
 148 sweet dream. End
 149 this ... okay? We can check that uh ... maybe some
 grammatical errors or

Pair 13: F-LRE (Grammar) (✓)

Excerpt 2. Pair 8, Billie and Jo, are focussing on the correct tense for the verb 'drive'. Jo tells Billie to take out the 'were' before 'driving' (Line 70), so this LRE was coded as incorrectly resolved.

- 64 J: Oh! While they were driving ... dri-ving
 65 B: Driving fast or?
 66 J: Uhh.. maybe!
 67 B: They should!
 68 J: Driving ...
 69 B: Ok never mind ... While they were driving ...
 70 J: Just driving
 71 B: Ok

Pair 8: F-LRE (Grammar) (X)

Excerpt 3 shows how Joanna and Cindy (Pair 11) had a spelling issue. Joanna asked Cindy how to spell McDonald's, but Cindy only sounded out the word. Instead of trying to write the word, Joanna suggests skipping it by using KFC instead. This LRE was coded as unresolved.

Excerpt 3. Unresolved LRE

- 280 J: 麥當勞怎麼拼?
(How do you spell McDonalds?)
- 281 C: Mac Donald m-a ... m-c ...
- 282 J: 算算算了我們跳過
(Forget it ... let's just skip it)
- (0.5)
- 283 C: KFC
- 284 J: 好 KFC Haha
(Ok)
- 285 C: Hahaha

Pair 11: L-LRE (Spelling) (?)

Patterns of interaction

The patterns of interaction were identified based on Storch's (2002) four patterns of interaction: collaborative; expert-novice; dominant-dominant; and dominant-passive. In this study, a fifth pattern known as cooperation was also included, as some learners may interpret collaborative writing differently (see Tan *et al.*, 2017). In Figure 1, cooperative patterns would fall into the same quadrant as the dominant-dominant pattern, as both patterns of interaction have a low mutuality level and high equality level. This meant that learners with a dominant-dominant patterns and cooperative patterns spend less time building on each other's ideas or suggestions (low engagement) and focused more on including their own ideas in the text (high contribution levels). The differences between these two patterns, however, are shown through the way learners negotiate before or during the writing stages. For example, cooperative pairs tend to distribute the workload equally, so that each person has the opportunity to contribute their own ideas in their allocated sections,

thus, spending less time listening to each other's ideas. In contrast, more disagreements occur during dominant-dominant interactions, as each person may try to take control of the writing task.

Similar to previous studies, three main patterns of interaction: collaborative, dominant-passive and expert-novice were identified in the fifteen transcripts (e.g., Storch & Aldosari, 2012). However, in addition to these three patterns, cooperative interactions were also found. Dominant-dominant patterns were not present cooperative interactions. The following excerpts show the four patterns of interaction found in the data.

Excerpt 4 shows a collaborative interaction pattern in which a pair listens to each other's ideas and suggestions. In this example, Joanna and Cindy both ask and answer questions, showing a high level of mutuality. Their ability to resolve each other's linguistic issues (Line 127, 130, 132, 135) shows a high level of equality as well.

Excerpt 4. Collaborative (Pair 11)

- 127 J: They were arrived the ... to the 喔還是 the?
(or)
- 128 C: 不用 to the arrive.
(Don't need)
- 129 J: Arrive 哪裡 ... 你要 arrive 哪裡 你要他們到哪裡,
或是他們到了就好
(*'Arrive' where ... Where do you want to arrive? Where do you want them to go, or just they arrive is ok?*)
- 130 C: 森林. 怎麼講?
(*Forest. How do you say that?*)
- 131 J: Forest f ... f-o-r-r-e-s-t 嗎?
- 132 C: 森林欸在森林野餐會怪怪的嗎?
(*Forest...Is it strange to go to the forest for a picnic?*)
- 133 J: 對阿. 可是這看起來就像森林不然就 park 公園 park.
(*Yeah. But it looks just like a forest. Or else, just park*)
- 134 C: 阿公園? 公園.
(*Ah, park? Park.*)
- 135 J: 對阿 滿像的啊. 好啊 arrive park 直接 park 不用 to

- the 喔 不用就 the
 (Yeah, looks quite similar ah. Ok ah. 'Arrive park' Just write 'park' directly, don't need 'to the' ah. Don't need. Just 'the park'.)
- 136 park.
 137 C: The park?
 138 J: 對.
 (That's right.)

Excerpt 5 shows an example of an expert-novice interaction in which Pete (expert) tries to encourage John (novice) to speak more English during the writing activity. He also tries to get John to engage with the task by asking him what certain things are in the picture.

Excerpt 5. Expert-Novice (Pair 10)

- 52 P: They opened their car and took out ...Took out 什麼?
 (what?)
- 53 J: Took out a ...竹籃子要怎麼說? What is the English name of 竹籃子?
 (How do you say Bamboo basket?) (Bamboo basket?)
- 54 P: Bamboo basket!
- 55 J: Bamboo?!
- 56 P: Mmm.
- 57 J: Bamboo 是抽煙的煙吧?
 (Bamboo is the smoke from smoking, isn't it?)
- 58 P: What is this?
- 59 J: 毛巾 towel!
 (Towel)
- 60 P: In English?

Excerpt 6 shows an example of a dominant-passive interaction in which Mary is more actively involved with the task and Shelly just agrees but does not provide suggestions.

Excerpt 6. Dominant-Passive (Pair 2)

- 119 M: Open the blanket sounds strange ... it sounds weird.

- 120 S: Yeah.
 121 M: Just throw? Throw out? Um ... after... he put the
 blanket on the
 122 beach ... can I say that?
 123 S: Yeah.
 124 M: Does it make sense?
 125 S: Yeah.

Excerpt 7 illustrates a cooperative interaction pattern. Belinda and Eric wrote separately and were responsible for their allocated sections, creating high levels of mutuality. This was shown through the pairs' use of first person 我 (*I*) in line 129 and second person *you* in line 128 when Belinda said that Eric had forgotten to mention the bag in his section. Moderate levels of equality, however, were present during their interactions, as they read each other's contributions.

Excerpt 7. Cooperative (Pair 7)

- 128 B: You didn't mention the bag...
 129 E: Oh, 我寫是這張
 (*But I was writing this one ...*)
 130 B: Oh! I lied then. Ok what she's doing now? What's she
 doing now

Interrater reliability

To ensure that the data was coded reliably, twenty per cent of the data was analysed again by a second year Masters of Applied Linguistics student, who was a native speaker of Mandarin Chinese. Six individual and three pair texts were coded in terms of CAF units and LREs). The percentage of agreement for the CAF units were high. The level for individual texts was 91.27% (DC, 84.6%; C, 95.29%; T, 84.06%; EFT, 93.33%; E, 98.85%), and the level for the collaborative texts was 93.89% (DC, 88.89%; C, 97.78%; T, 97.3%; EFT, 89.47%; E, 96%). For the LRES, the percentage of agreement scripts was moderately high (87.86%) (L-LRE, 88.89%; M-LRE, 93.75% and F-LRE, 80.95%). As the percentages of agreement were all within the range of 75% to 90% (Hartmann, 1977), this demonstrated an acceptable level of agreement.

RESULTS

This section reports the findings in terms of CAF measures (RQ1), and LREs and patterns of interaction found in the pair talk (RQ2).

Research Question 1

The first research question compares the quality of individual and collaborative texts in terms of complexity (C/T; DC/T), accuracy (EFT/T; E/T) and fluency (W/EFT) values. As multiple pair sample t-tests were conducted, a Bonferroni correction was done, and the alpha level was set at .008.

Complexity

The descriptive statistics for individual and collaborative text complexity in Table 4 show that individuals produced more clauses per T-unit (C/T) than pairs, but a very similar numbers of dependent clauses per T-unit (DC/T).

TABLE 4
Complexity (n=30)

	CI 95%					
	M	SD	Min.	Max.	Lower	Upper
C/T						
Individual	1.33	0.2	1.00	1.9	1.25	1.41
Collaborative	0.92	0.24	0.57	1.7	0.83	1.01
DC/T						
Individual	0.36	0.24	0.00	1.25	0.27	0.45
Collaborative	0.34	0.13	0.00	0.6	0.29	0.39

The paired- sample t-tests showed that number of clauses per T-unit was significantly higher in the individual texts than in the pair texts ($p < 0.008$), $t(29) = -5.89$, $p = 0.01$, $r = 0.1$. According to the aforementioned guidelines, the effect size however, was only small. The difference between DC/T for individual and pair texts was not statistically significant ($p > 0.008$).

Accuracy

Table 5 shows the descriptive statistics for both individual and collaborative texts. The paired-sample t-tests indicate that for both measures pairs' text accuracy was significantly higher than for the individual texts: EFT/T, $t(29)=3.85$, $p=0.001$, $r=0.2$, and E/T, $t(29)=-4.07$, $p<0.001$, $r=0.3$. The effect sizes varied; E/T showed a medium effect size, whereas EFT/T showed a small effect size.

TABLE 5
Accuracy (n=30)

	M	SD	Min.	Max.	CI 95%	
					Lower	Upper
EFT/T						
Individual	0.24	0.22	0	0.83	0.16	0.33
Collaborative	0.44	0.23	0.07	0.8	0.35	0.53
E/T						
Individual	1.55	0.7	0.28	3.5	1.29	1.81
Collaborative	0.99	0.52	0.33	2.2	0.8	1.18

Fluency

Table 6 shows that the average number of words (W) and the number of words per error-free t-units (W/EFT) produced by pairs and individuals. The paired-sample t-tests showed that there were significant difference in both measures, with pairs producing more fluent texts than individuals ($p<0.008$): W, $t(29)=2.95$, $p<0.006$, $r=0.4$. and W/EFT, $t(29)=2.66$, $p<0.001$, $r=0.1$. Similar to the text accuracy levels, the effect sizes also varied; W showed a medium effect size, whereas W/EFT showed a small effect size.

TABLE 6
Fluency (n=30)

	M	SD	Min.	Max.	CI 95%	
					Lower	Upper
W						
Individual	118.7	33.2	39	179	106.09	131.31
Collaborative	142.87	45.15	69	263	121.8	164.02
W/EFT						
Individual	7.06	5.36	0	18	5.06	9.06
Collaborative	9.9	2.62	5.33	14.5	8.92	10.88

The smaller standard deviation for fluency compared to accuracy indicates that the writers did not vary as much in terms of fluency as they did for accuracy. Overall, the findings show that students were able produce better quality texts when they worked collaboratively in regard to accuracy and fluency. However, in contrast, for complexity, pairs produced less complex texts on one of the measures, and there was no difference for the other measure.

Research Question 2

Table 7 illustrates pairs' patterns of interaction, along with the average time spent on the task, the average number of turns taken, and the average number of turns containing LRES. The average number of turns provides of the amount of discussion engaged in by pairs during the collaborative writing activity. The percentages of LREs in pair talk were based on the average number of LREs over the total number of turns; this determined the proportion of time spent on language-related issues during the writing task.

TABLE 7

Average time taken, turns and LREs for patterns of interaction

	Pairs	Average time spent on task (mins)	Average number of Total Turns	Average number of LRE turns	The percentage of LREs turns in pair talk (LRE/Turns in %)
Collaborative	7	27.04	252	134.6	51.58
Dominant-Passive	4	19.69	184	102.25	55.64
Expert-Novice	3	29.88	201	73.75	36.75
Cooperative	1	38.4	162	50	30.86
All pairs	15	26	217	113	52

The findings reveal that pairs who worked collaboratively produced the highest number of turns and LRE turns. Dominant-passive pairs had the second highest number of LRE turns, although they had fewer turns than did Expert-Novices. Expert-novice pairs spent a long time on the task but produced the smallest number of LRE turns. There was only one pair who worked in a cooperative pattern; this pair engaged in the least amount of discussion (based on turns and LRE turns), but spent the largest amount of time working on the task (38.4 minutes).

Table 8 shows the findings for the LREs and interaction patterns produced during the collaborative writing task. Even though the task instructions focused on past-tense usage, many of the LREs produced were lexis-focused, with students focussing mainly on discussing word definitions and choices. This was followed by a focus on mechanics and then on form, although the difference between the focus on mechanics and form was only small. Students resolved 246 out of the 333 LREs (73.87%) correctly without teacher-assistance and relatively few LREs were resolved incorrectly (9.01%) or remained unresolved (17.12%). This highlights how collaborative writing tasks can provide students with opportunities to discuss and solve many linguistic issues without teacher assistance.

TABLE 8
LREs (n=15)

	✓	X	?	Total
Lexis-focused	121 (36.34%)	9 (2.7%)	40 (12.01%)	170 (51.05%)
Mechanics-focused	70 (21.02%)	6 (1.8%)	12 (3.6%)	88 (26.43%)
Form-focused	55 (16.52%)	15 (4.51%)	5 (1.5%)	75 (22.52%)
Total	246 (73.87%)	30 (9.01%)	57 (17.12%)	333

Notes: ✓=correctly resolved; X=incorrectly resolved; ?=unresolved

The results presented in Table 7 and Table 8 provide an impression of how pairs' patterns of interaction influenced the amount of discussion which occurred. While according to Table 7 and Table 8 only one pair showed cooperative patterns of interaction in which the pair focused on their own sections (and discussed during the swap over), a closer qualitative examination of the data revealed that some pairs also divided the work load, although in a different way. The effect that task division had on pairs writing depended strongly on their levels of engagement and contributions to the task. Unlike Pair 7, in the other pairs one person discussed (or dictated) what to write, while the other wrote. When the pairs acknowledged each other's suggestions and opinions while they were writing, there was a higher level of engagement and contribution. Role changes allowed passive students to contribute to the writing task (Pair 2), while encouraging novice students to reach their potential writing level with assisted help (Pair 9).

Pair 2 (dominant-passive pair) provides an example of how a shift in roles can affect the patterns of interaction. In the first half of the task, Mary (dominant) wrote down the ideas that Shelley (passive) provided (or dictated). During this stage of the interaction, Mary still actively contributed to the task:

Excerpt 8. Pair 2

- 55 S: He went to a beach
56 M: Hmm?
57 S: A beach
58 M: Uh ... Can I said he went to a islands right?
59 S: Went to fantastic island

This role as dictator gave Shelley a chance to participate in the task. When it was Shelley's turn to write, she contributed more to the text than to the discussions. In other words, she did not question or provide as many suggestions as Mary had done; instead, she made her own changes to the actual text; this shows a high level of equality, as Shelley tries to contribute her own ideas to the text.

Excerpt 9. Shows an example of Shelly and Mary's dominant-passive interaction pattern:

- 183 M: Even he has some
184 S: He has ...
185 M: Some of slave,
186 S: Huh?
187 M: Help him ... help him ... massage him

Shelley is seen as the passive student during these discussions, as she only repeated utterances (line 184) and used a one-word exclamation (line 186). In the text however, Shelley wrote: *Even he had some of the service, which helped massage*. This sentence shows that Shelley used *of the service* instead of *slave* after she said *huh* (line 186) during the discussion with Mary. Even though the phrase Shelley used was grammatically incorrect, it shows how Shelley was still able to contribute a little more to the task even though she was very passive during the discussions.

DISCUSSION

The findings illustrate the impact collaborative writing can have on a pen-and-paper writing talk in an EFL classroom context; the analysed pair talk and texts show how collaborative writing tasks can

encourage learners to pool together and apply their knowledge to the process of writing a text.

Peer interaction and text quality

In line with many previous studies, the participants were able to produce more accurate texts when they co-constructed with a partner (Dobao, 2012; Wigglesworth & Storch, 2009). The peer discussions also encouraged learners to build on each other's ideas to create longer texts. The co-construction allowed pairs to engage in meaningful conversations about lexical choices. In contrast to previous studies (Storch, 2008), the findings showed that students spent more time on resolving vocabulary than discussing spelling or grammar. The learners' word choices at times had an academic tone through their use of connectives (Pair 2, Pair 5) and/or low-frequency words (Pair 9), even though this is not very appropriate for story-writing picture ordering tasks. These choices could illustrate the influence of the exam-oriented classroom, which prioritizes more formal academic language use. Genre-awareness is an issue that classroom teachers need to address, as it is important that students are aware of how different written genres can influence vocabulary choices.

In contrast, the findings showed that learners in collaborative pairs produced less complex texts than the individuals in terms of clauses per T-unit. However, there were no significant differences shown for the use of dependent clauses. Thus, it seems collaborative writing had a positive influence on accuracy and fluency, but not on complexity. The findings LRES revealed that the learners focused more on vocabulary and spelling than on structural aspects of writing. This could explain why these collaborative writing activities did not have a significant effect on text complexity.

The impact of interactive patterns

Four main patterns of dyadic interaction were found in this study: collaborative, dominant-passive, expert-novice and co-operative. Like previous studies, pairs worked in mainly collaborative, dominant-passive and expert-novice patterns, with the collaborative pattern being the most common pattern (Storch & Aldosari, 2013; Watanabe & Swain, 2007). Collaborative pairs were able to produce more LREs

and resolve linguistic issues in a co-constructed manner. Their equal levels of contribution and engagement gave them the opportunity to pool their existing linguistic resources and use these as a tool to resolve issues effectively (Storch, 2002).

Notably, task distribution was a common approach pairs used to complete the writing task. The findings revealed that this approach affected the pairs differently depending on the nature of their interactions. For example, when the pairs wrote and discussed simultaneously, their constant change of roles gave them the opportunity to engage further with the task (e.g., Pair 2 and 5), thus balancing the levels of equality and mutuality to create a collaborative relationship. However, sometimes when pairs separated workloads, a stronger emphasis was made on the individual's responsibility for completing the allocated sections (e.g., Pair 7); this created a cooperative pattern of interaction, where students tend to treat their sections as solitary acts. This co-operative pattern is a pattern that very few studies have identified. Tan *et al.* (2010) only found cooperative patterns through CMC, but not in face-to-face interactions.

Each pattern of interaction provided opportunities for the writers to negotiate meaning. When students worked together, collaboration provided individuals with the chance to extend themselves with assistance from their partner. The pedagogical implications of this study are that paper-based collaborative writing can be effective for learners who are used to exam-oriented EFL classrooms. The interactive process involved in collaborative writing can encourage students to revise, recall and apply language forms and functions effectively. More attention, however, is needed to understanding students' attitudes towards collaborative writing and how they initially approach writing tasks, through planning, negotiating meaning, dividing up the task and so forth.

It is important to stress that students' attitudes towards and approaches to collaborative writing can vary depending on sociocultural differences, language proficiency and task type. The current study has only examined lower linguistic aspects of writing through measures of CAF. More research on higher-level conceptual aspects of writing, as well as different writing genres, and different

classroom contexts could provide a more holistic view of how paper-based collaborative writing activities influence writing.

THE AUTHOR

Annita Stell is a PhD student in Applied Linguistics at the University of Queensland, Australia. Her research interests include peer interaction, mediation and collaborative writing in second language classrooms. Her specific interest lies in how pairs' writing strategies impact their individual L2 writing development. She currently teaches academic and creative writing to ESL learners in Australia.

REFERENCES

- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*(0), 42-65.
- Butler, Y.G. (2015). English language education among young learners in East Asia: A review of current research (2004-2014). *Language Teaching, 48*(3), 303-342.
- Cambridge English Language Assessment (2015). About the Common European Framework of Reference for Languages (CEFR). *International language standards explained*. Retrieved from www.cambridgeenglish.org/exams/cefr/.
- Chou, C.P., & Yuan, J.K.S. (2011). Buxiban in Taiwan. *The Newsletter: International Institute for Asian Studies, 56*.
- Chu, H-Y., & Yeh, H-N. (2017). English benchmark policy for graduation in Taiwan's higher education: Investigation and reflection. *Journal of Language Teaching and Research, 8*(6), 1063-1072.
- Chung, I.F. (2013). Crammed to learn English: What are learners' motivation and approach? *Asia-Pacific Education Researcher, 22*(4), 585-592.
- Cumming, A. (2012). Goal theory and second-language writing development, two ways. In R. Manchon (Ed.), *L2 writing development: Multiple perspectives* (pp.135-164). Berlin: De Gruyter Mouton.

- Dobao, A.F. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*(1), 40-58.
- Donato, R. (1994). Collective scaffolding in second language learning. In J.P. Lantolf & G. Appel (Eds), *Vygostskian approaches to second language research* (pp.33-56). Norwood, NJ: Ablex.
- Educational Testing Service. (2012). Scores Overview. *The TOEFL family of Assessments*. Retrieved from etsglobal.org/Global/Eng/Tests-Preparation/The-TOEFL-Family-of-Assessments/TOEFL-ITP-Tests/Scores-Overview.
- Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology, 14*(3), 51-71.
- Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. In A. Housen, F. Kuiken & I. Vedder (Eds), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp.247-276). Amsterdam: John Benjamins.
- Glendinning, E., & Howard, R. (2003). Lotus ScreenCam as an aid to investigating student writing. *Computer Assisted Language Learning, 16*(1), 31-46.
- Haneda, M. (2007). Modes of engagement in foreign language writing: An activity theoretical perspective. *Canadian Modern Language Review, 64*(2), 297-327.
- Hartmann, D.P. (1977). Considerations in the choice of interobserver reliability measures. *Journal of Applied Behavior Analysis, 10*, 103-116.
- Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.
- Kowal, I. (2014). Fluency in second language writing: a developmental perspective. *Studia Linguistica Universitatis Iagellonicae Cracoviensis, 131*, 229-246.
- Lantolf, J.P., & Poehner, M.E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11-33.

- Li, M., & Storch, N. (2017). Second language writing in the age of CMC: Affordances, multimodality, and collaboration. *Journal of Second Language Writing, 36*, 1-5.
- Liao, C., Chang, W., & Chan, T. (2018). The effects of participation, performance, and interest in a game-based writing environment. *Journal of Computer Assisted Learning, 34*(3), 211-222.
- Ministry of Education Republic of China (2018). *Ministry of Education Objectives for 2018 (January to December). Plans and objectives*. Retrieved from english.moe.gov.tw/cp-9-17161-B6F6B-1.html.
- Mirzaei, A., & Eslami, Z.R. (2015). ZPD-activated languaging and collaborative L2 writing. *Educational Psychology, 35*(1), 5-25.
- 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.
- Pan, Y-C. (2015). Test impact: English certification exit requirements in Taiwan. *TEFLIN Journal, 20*(2), 119-139.
- Pan, Y.C., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching, 9*(1), 108-122.
- Philp, J., Adams, R.J., & Iwashita, N. (2014). *Peer interaction and second language learning*. New York: Routledge.
- Sajedi, S.P. (2014). Collaborative summary writing and EFL students' L2 development. *Procedia-Social and Behavioral Sciences, 98*, 1650-1657.
- Shohamy, E. (2017). Critical language testing. In E. Shohamy, I. Or & S. May (Eds), *Language testing and assessment* (pp.441-454). New York: Springer.
- Stell, A. (2015). *The effectiveness of implementing collaborative writing as an EFL classroom activity for L2 development*. Masters thesis, University of Queensland, Brisbane, Australia.
- Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research, 5*(1), 29-53.
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153-173.

- Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness, 17*(2), 95-114.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual matters.
- Storch, N., & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research, 17*(1), 31-48.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J.P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp.97-114). Oxford: Oxford University Press.
- Swain, M. (2013). Cognitive and affective enhancement among older adults: The role of languaging. *Australian Review of Applied Linguistics, 36*(1), 4-19.
- Tan, L.L., Wigglesworth, G., & Storch, N. (2010). Pair interactions and mode of communication. *Australian Review of Applied Linguistics, 33*(3), 27.1-27.24.
- Test of English as a Foreign Language (1994). Test 1. Section 2: *Structure and written expression* (pp.14-20). Princeton, NJ: Educational Testing Service.
- van Compernelle, R.A., & Williams, L. (2013). *Sociocultural theory and second language pedagogy*. London: Sage.
- Vygotsky, L. (1987). *Mind in society: The development of higher psychological processes*. Cambridge, MA: 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.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*(2), 121-142.
- Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and instruction, 10*(1), 37-53.

- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H-Y. (1998). *Second language development in writing: Measures of fluency, accuracy & complexity*. Honolulu: University of Hawaii Press.
- Wu, T.Y. (2011). Learning English in Taiwan's elementary schools. *Journal of Futures Studies*, 16(2), 35-46.

APPENDIX A: SCORE CONVERSION**Score conversion**

IELTS bands	CEFR Levels	Structure and Written Expression	Raw TOEFL scores
8-7	C1 ***	64	37
6.5-5.5	B2 **	53	29
5-4	B1 **	43	18
	A2 *	32	10

* Basic user (A1-A2) **Independent user (B1-B2) ***Proficient user (C1-C2)

(Adapted from Cambridge English Language Assessment, 2015; Educational Testing Service, 2012; Test of English as a Foreign Language, 1994)


APPENDIX B: PARTICIPANTS

Pair	Student	Raw Scores	Converted scores	CEFR
1	Emma	14	38	A2
	Anny	8	27	A1
2	Mary	15	40	A2
	Shelly	27	51	B1
3	Andy	34	58	B1
	Crystal	23	47	B1
4	Ally	21	46	B1
	Lily	25	49	B1
5	Zack	20	45	B1
	Vincent	25	49	B1
6	Samantha	30	54	B1
	Tom	35	59	B1
7	Belinda	26	50	B1
	Eric	21	46	B1
8	Jo	29	53	B2
	Billie	24	48	B1
9	Ben	24	48	B1
	Jimmy	23	47	B1
10	John	35	59	B2
	Pete	29	53	B2

11	Joanna	22	47	B1
	Cindy	15	40	A2
12	Linda	17	42	A2
	Becky	15	40	A2
13	Amy	26	50	B1
	Jenny	33	57	B2
14	Julie	25	49	B1
	Catherine	20	45	B1
15	Milly	10	32	A2
	Andy	34	58	B2

APPENDIX C: TASK INSTRUCTIONS

- Pair writing task instructions

NAMES: _____ _____
<hr/> •PAIR WORK • Task: Rearrange the comic in the correct order, and write a paragraph about it in the space below (Use <u>past tense</u>). Duration: 30 minutes Record the activity with your mobile/cell phone 

- Individual writing task instructions

NAME: _____ _____
<hr/> •INDIVIDUAL ACTIVITY • Task: Rearrange the comic in the correct order, and write a paragraph about it in the space below (Use <u>past tense</u>). Duration: 20 minutes

- An example of one of the four Garfield comics used for the study

