DOI: 10.30050/CLLT.202206_2(1).0003

Task Repetition with Peer Review as Corrective Feedback for L2 Writing in Mandarin Chinese

Miao-fen Tseng*

Ziyi Geng

University of Virginia

Abstract

Studies of task repetition have yielded fruitful results in terms of improvement of L2 oral performance (Bygate, 2001). However, the effects of task repetition on L2 writing have not yet been explored in writing in Mandarin Chinese. Drawing on Bygate's (1999) argument that task repetition involves the framing and reframing of language in meaningful and communicative contexts, this study investigates the impact of task repetition on L2 expository writing in Mandarin Chinese, with peer review functioning as corrective feedback and intervention. Twenty-two learners enrolled in a third-year Chinese language course were invited to construct and reconstruct their writing discourse through peer review. The peer review process featured a combination of computer-assisted and virtual face-to-face interaction. After repeating the writing task, learners completed a survey on the pedagogical value and design of peer review. Data analysis included descriptive analysis and statistical analysis through t-tests. The results of task repetition in writing revealed that the repeated tasks yielded different degrees of improvement in accuracy, fluency, and

* Miao-fen Tseng is the corresponding author.

complexity. The survey results also reaffirmed the value of task repetition through peer review as a means for corrective feedback. Pedagogical implications, limitations of the study, and future research are outlined at the end of the paper.

Keywords: task repetition; peer review; L2 writing; complexity, accuracy, and fluency; task-based language teaching

1. Introduction

The number of studies in task-based language teaching (TBLT) has skyrocketed in second language acquisition and foreign language pedagogy over the past three decades. With versatile models and multi-dynamic interpretations, the theoretical underpinnings of TBLT continues to evolve, while maintaining focus on student-centeredness and authentic communication.

A task is a dynamic and powerful construct essential to TBLT. It provides a meaningful, communicative, and authentic context for learners to shape and reshape meanings through the process of meaning negotiation. Although there is no single agreed-upon definition of a task, characteristics can be identified, drawing upon several common features according to Skehan's (1998) and Ellis's (2003) studies. A task places meaning and communication at the heart of the curriculum. Learners immerse themselves in "real-life" situations and use their own language resources to communicate meaning, rather than relying on a set of predetermined rules or patterns partially or fully controlled by the instructor. Clearly defined learning objectives are pivotal for assessments, as they directly impact the learning outcomes and task completion.

Tasks are classified according to criteria including authenticity (Richards, 2001), structure-focused versus meaning-focused (Ellis, 2004; Skehan, 1998), pedagogical objectives (Willis & Willis, 2007), and interactional patterns (Pica, Kanagy, & Falodum, 1993). Solidly grounded in TBLT theoretical frameworks, Tseng (2014, 2018, 2020) has created a plethora of different types of tasks according to proficiency level, learning objective, topic, and task type, with task-specific rubrics for performance-based assessments in teaching Mandarin Chinese as a foreign language. The wide array of tasks can be used intact or modified for TBLT curricula (Long, 1988; Willis, 1993; Long & Crookes, 1992) or task-supported language teaching (TSLT) curricula (Skehan, 1998; Ellis, 2003; Littlewood, 2007; Samuda & Bygate, 2008). While TBLT places tasks at the heart of curricular design, TSLT treats tasks as supplementary components of a PPP (present, practice, and produce) approach or other methodology, such that they are typically implemented as exit tasks for summative assessment. Apparently, TSLT yields much more

flexibility, applicability, and practicability for the design of language courses and programs and the fulfillment of local needs in different teaching contexts.

Numerous empirical studies have tested different dimensions of TBLT and TSLT, including effects of task variables on learners' language performance, motivation, and inter-language development (Robinson, 2011); independent measures of task complexity (Revesz, Sachs, & Hama, 2012); grammar teaching through focus-on-form instruction (Hossein & Fotos, 2011; Fotos & Nassaji, 2007), and teachers' perception and program design (Tseng 2017, 2018, 2019; Van den Branden, Bygate, & Norris, 2009). Single tasks have received far more attention than task repetition. Although the concept of task repetition was mentioned in the original three-stage framework of TBLT—pre-task, core-task, and post-task (for example, Willis, 1996)—it has not been as frequently discussed as single tasks until very recently (for example, Bygate, 2018).

Inspired by Khezrlou's (2020) study of task repetition, which found positive effects on writing improvement through error correction, this study takes task repetition as a starting point to examine L2 writing performance in Mandarin Chinese, with peer review as a means of intervention for corrective feedback between the two task instances. This experimental design is the very first attempt to examine complexity, accuracy, and fluency of repeated tasks in L2 writing and to explore the pedagogical values and effective design of task repetition in teaching and learning writing in Mandarin Chinese.

2. Literature review

2.1 Task repetition

Inspired by Levelt (1989), Bygate (1996, 2001) related Levelt's speech production model to the positive psychological effects of task repetition on L2 learning. In Levelt's model, speech is produced through three interrelated phases: conceptualization, formulation, and articulation. The three-staged process suggests that while performing oral tasks, learners first concentrate on content generation. Repeating a task helps them to conceptualize content, thereby enabling them to meticulously determine language choices

and monitor language outputs on subsequent iterations. This alters learners' focused attention from content to accurate and appropriate formulations, such that they move from meaning-oriented to form-oriented production (Bygate, 1996). Task repetition does not engage invite learners to do the "same" thing, but rather to work differently on the same material (Bygate & Samuda, 2005). It can be reasonably argued that trade-off effects may decrease as repetitions increase.

General speaking, there are three ways to operationalize task repetition: 1) repeat the exact same task (i.e., exact task repetition); 2) repeat the same procedure with different content (i.e., procedure repetition); and 3) repeat the content of the task with a different procedure (i.e., content repetition) (Kim et al., 2018). Task repetition is the same as language reproduction or the precise repetition of the language used. A task is repeated with "a given configuration of purposes, and a set of content information" that allows additions, omissions, or "possible substitutions of language or moves" (Bygate, 2018). Larsen-Freeman (2018) recommends task *iteration* in lieu of task *repetition* according to Complex Dynamic Systems Theory (CDST). She identifies three key notions of task iteration according to CDST: learners' agency and uniqueness, learning not limited to reproduction, and transformation rather than transfer.

To date, empirical studies have reported some effects of task repetition on the development of L2 oral development in adults learning English as a foreign language. Plough & Gass (1993) documented that immediate repetition of a task revealed learners' willingness to clarify form through meaning negotiation, but that some students were demotivated in the repeated task. In a case study in which a story narrative task was repeated after a three-day interval, significant improvement was found in four aspects: accuracy, idiomaticity, lexis, and self-correction (Bygate, 1996). Similar effects were found in a series of six repeated case studies by Lynch & Maclean (2000, 2001). In a poster carousel activity, learners repeated summaries of the content of the poster over different visits; as the visits progressed, they improved in pronunciation, grammatical accuracy, and lexical expressions. In Bygate's (2001) study, a large-scale repetition of a single task after a 10-week interval showed that while fluency and complexity increased, accuracy was unaffected in learners' oral story narration. Further analyzing the 2001 data sets, Bygate &

Samuda (2005) found increased elaboration in the use of adverbials indicative of place, time, manner, cause, and purpose. Task repetition in response to a video-based monologue story narration has shown an impact on speech performance at the discourse level, in addition to effects on accuracy, fluency, and complexity (Wang & Chen, 2018). A task that was repeated with five encounters intact, each in a two-week interval, seemed to have effects on the same constellation of task materials, but not on different content materials (Kim et al., 2018).

Task repetition yields several advantages. Task-as-process can free up learners' attentional resources (Bygate, 2001; Bygate & Samuda, 2005). Within-task repetition promotes automaticity in elicited formulaic speech (Gatbonton & Segalowitz, 2005). Task-type repetition enhances writing performance (Nitta & Baba, 2014). Repeated practice in the procedure of task repetition contributes to the reconstruction of language rules and usage (de Jong & Perfetti, 2011). Learners' familiarity with a task potentially increases the number of interactional exchanges, such as requests for clarification, reformulation, and confirmation, and language-related episodes (Kim & Ventura-Tracy, 2013). Bygate (2018) poses that task repetition has storage and retrieval strengths in multiple ways. After looking into studies on task repetition, Bygate makes three provisional conclusions: First, significant differences were found in one or more aspects of performance in fluency, accuracy, and complexity between iterations. Second, iterations in various aspects of language are likely to arise, to some extent irrespective of proficiency level. Third, although learners' iterations are likely to improve, it cannot be predicted whether this will predominantly occur in fluency, accuracy, and complexity.

2.2 Cognition hypothesis and trade-off hypothesis

According to Robinson's Cognition Hypothesis (2001, 2003), tasks can be sequenced by cognitive complexity, as defined through tasks' resource-directing variables, pertaining to the use of linguistic forms, and resource-dispersing variables, such as pre-task planning. The hypothesis poses three descriptors or indicators for L2 language performance: complexity, accuracy, and fluency (hereafter CAF). Robinson claims that learners can simultaneously access multiple and multidimensional attentional resources to

operationalize and manipulate task complexity by increasing the cognitive demands of a task, which leads to simultaneous improvement in complexity and accuracy.

Due to limitations on capacity and working memory in processing the target language, during a task, learners must allocate and reallocate their attentional resources in the learning processes, such as selecting input, processing information, revisiting the interlanguage system, and finalizing output, to name a few. This brings about competition among CAF for attentional resources, leading to a "trade-off" hypothesis (Skehan, 2009), known previously as the Limited Attentional Capacity Model (Skehan, 1998; Skehan & Foster, 2001). According to the hypothesis, one aspect is achieved at the expense of the remaining two, likely because learners consciously or subconsciously focus on one dimension to the detriment of the other two. In preparing or planning for an oral task that involves expressing or negotiating meaning, fluency is normally prioritized over accuracy and complexity. Even if attentional resources are devoted to the accuracy of form, there is still competing tension between fluency and complexity. This competition among CAF has incentivized researchers and practitioners to explore the impact of task difficulty on language output through the analysis of CAF as performance or progress descriptors or indicators underlying proficiency.

In the past three decades, a wealth of studies examined task complexity and its impact on CAF in L2 production, focusing on speaking more than writing. Peter Skehan has frequently worked with Pamela Foster to examine how various factors affect learners' L2 production, mainly in oral tasks. Structural complexity and lexical complexity are now analyzed separately by Skehan, resulting in the terminology of CALF in the most current literature with the addition of lexical complexity. Johnson's (2017) quantitative meta-analysis verifies that increased task complexity in resource-directing and resource-dispersing features impacts CALF in written L2 production. Although no clear evidence was found to support the cognition hypothesis (Robinson, 2001, 2003, 2005, 2011), task complexity may be beneficial for attentional resources to promote the formulation and monitoring systems in L2 writing (Kellogg, Whiteford, Turner, Cahill, & Mertens, 2013).

2.3 Complexity, accuracy, and fluency

Many researchers and practitioners assume that the constructs of language performance are multifaceted and multi-componential in nature and that they can be adequately captured by CAF (Skehan, 1998; Ellis, 2003, 2008). Despite extensive research in CAF, none of these three constructs is uncontroversial in its operational definition. Accuracy is probably the most transparent and consistent construct of the triad (Hammerly, 1991). In general, there is less debate in assessing the accuracy of L2 writing than in assessing L2 oral performance, in terms of accounts of errors in comparison with different linguistic aspects of standard and nonstandard usage among native speakers (James, 1998; Ellis, 2008).

According to Johnson's (2017) research synthesis and meta-analysis, accuracy, strikingly, is comparatively rarely investigated in L2 writing. In a few studies, the number of errors in T-units or error-free units was used to analyze L2 writing accuracy. Fluency measures appeared to be even narrower than accuracy measures. The total number of words produced per minute was considered an indication of writing fluency within time constraints. In examining complexity in L2 writing, clauses per T-unit were the metric most frequently used; the second-most used metric was the mean length of per T-unit. Narratives were the most frequently studied genre. The analysis of expository and argumentative essays tended to focus on phrasal and clausal complexity. Lexical complexity was examined in terms of lexical diversity, sophistication, and/or lexical density. The majority of studies have relied on metrics of syntactic complexity associated with complex forms more typical of oral language production (Biber & Conrad, 2009; Biber & Gray, 2010; Biber, Gray, & Poonpon, 2011, 2013).

2.4 Peer review

As foreign language teaching undergoes a paradigm shift from teacher-controlled to student-centered pedagogy, L2 writing is no longer characterized merely as a linguistic product. As Kern (2000) wrote, it is "an integrative, student-centered approach . . . that also attends to the interdependencies among textual products, cognitive processes, and sociocultural factors" (p. 185). Taking into account the cognitive and sociocultural aspects of learning encourages learners' agency, self-regulation, and autonomy. In response to this

trend, peer review in corrective feedback is currently one of the theory-endorsed best practices in L2 writing.

The theoretical framework of peer review in L2 writing is grounded in various frameworks, including the process writing theory (Hayes & Flower, 1980), the sociocultural theory (Vygotsky, 1978), and the interactional theory (Swain & Lapkin, 1998, 2002). Peer review has long been recommended as a way to facilitate language learners' writing process. In L2 writing, it has been proven to motivate language learners to actively produce corrective feedback, instead of relying only on the instructor to do so, which may significantly enhance learner agency and autonomy in L2 writing and language learning (Hu, 2005; Hyland & Hyland, 2006). Peer review can encourage learners to actively examine their writing; search for potential problems, errors, and areas for improvement; and address errors through peer modifications or self-modifications based on peer feedback (Levi Altstaedter, 2016; Yang, 2011). It has also proven effective in engaging learners in scaffolding or assisting each other within the Zone of Proximal Development (ZPD) to edit and improve their writing (De Guerrero & Villamil, 2000; Hanjani & Li, 2014). Therefore, the peer review process can help learners to enhance the quality of their L2 writing through peer interaction.

The ways in which peer review is conducted vary greatly in both process and format. There currently exist two major modes of peer review in L2 writing: face-to-face peer review (FFPR) and computer-assisted peer review (CAPR). When employed alone, both FFPR and CAPR have advantages and disadvantages. It is reported that FFPR can better engage learners to seek and provide suggestions in text revision and evaluation through synchronous communication. In FFPR, learners not only can give each other feedback on writing but are also able to clarify ideas and exchange opinions. Therefore, learners engaging in FFPR were generally reported as benefiting from live peer interaction and better able to verbalize questions or express ideas while receiving immediate feedback from their partners (Ho & Savignon, 2007). CAPR is reported to offer benefits that FFPR does not, including efficiency in typing feedback; learners feeling less pressure than when pointing out each other's errors in FFPR; saving time in regular classes (Ho & Savignon, 2007); and learners paying more attention to language forms and structures (Saeed et al.,

2018). Researchers recommend the combined use of FFPR and CAPR, so as to take advantage of the merits of both modes (Ho & Savignon, 2007; Saeed et. al., 2018).

Any peer review process requires advance training and preparation. To foster peer interaction and make it as meaningful and productive as possible requires careful planning and preparation, with explicit instruction. Hence, learners should know what to do in peer review and how to do it well through hands-on experience coupled with modeling. Understanding teachers' expectations of their performance and working strategies helps learners make their comments focused and revision-oriented in both local and global textual revisions (Saeed et al., 2018). In studies comparing coached to uncoached groups of learners who then engaged in peer review, the coached groups were found to engage more extensively and actively than the uncoached groups (Zhu, 1995; McGroarty & Zhu, 1997) and, significantly, generated more comments in the interactive process (Min, 2005).

Currently, research into the peer review process in EFL/ESL settings has been fruitful, but very little has been done in the context of Chinese language teaching, especially in terms of peer review's effectiveness when combined with task repetition. The effects of peer review in task repetition on L2 writing have yet to be explored in teaching Chinese as a foreign language. This need serves as the momentum for the present study.

3. Research Questions

In the framework of task repetition, multiple tasks are carried out in a sequence. This creates opportunities for learners to revisit and fine-tune what they produced previously, both structurally and functionally, ultimately improving their language performance. Recognizing the pedagogical value of task repetition, this study generates two hypotheses. First, task repetition is beneficial for the improvement of writing performance in complexity, accuracy, and fluency. Second, peer review as a means of corrective feedback is helpful for enhancing writing performance in the repeated task. Specifically, the study seeks answers to the following four research questions:

1. Did task repetition with peer review as corrective feedback help to improve complexity in L2 writing in Mandarin Chinese? If yes, to what extent?
2. Did task repetition with peer review as corrective feedback help to improve accuracy in L2 writing in Mandarin Chinese? If yes, to what extent?
3. Did task repetition with peer review as corrective feedback help to improve fluency in L2 writing in Mandarin Chinese? If yes, to what extent?
4. What were learners' input and comments on the peer review process as a means of corrective feedback (intervention) between the two tasks?

This study takes CAF, but not CALF to analyze L2 writing; some language features in Mandarin Chinese differ from those in teaching English as a second or foreign langauge. Adjusting the categorization of data analysis is not unusual and indeed necessary in data-driven empirical studies. Furthermore, this study is the first attempt to examine the impact of task repetition on writing performance in teaching Chinese as a foreign language. Although findings are preliminary, they have potential to shed light on our understanding of repeated tasks, strategic design of peer review processes, and factors contributing to productive corrective feedback on writing performance that are less or not yet studied in repeated tasks.

4. Research Method

To answer the above four research questions, this study examines pre-advanced learners' writing performance in Mandarin Chinese. Specifically, it focuses on the results, in terms of CAF, of a repeated writing task: the production of an expository essay. The task repetition process was mediated through peer review. A combination of descriptive and statistical analysis through t-tests was employed to complement and strengthen the analysis of writing performance. An end-of-semester survey on a 7-point Likert scale, along with an open-ended question, was administered to gather learners' assessment of the design and effect of peer review as a means of corrective feedback.

This section describes the context in which learners studied and performed their writing, including how task cycle and task repetition were implemented in the curriculum, who the participants were, and what learners did to follow the peer review process.

4.1 Context

The study took place in a third-year Chinese language course at a public university in the United States. The course was taught purely online, featuring technology-mediated and task-based language teaching that places authenticity and tasks at the heart of the curriculum. It also aligns with Skehan's (1998) recommendations for predetermining target structures in content and material development. In the course, learners participated in a virtual US–China exchange program in which they embarked on a journey, landed at the Beijing airport, visited Chinese universities, explored historical sites and relics, and immersed themselves in different cities and areas. In the linguistically and culturally rich journey, learners experienced the multifaceted modernity, traditions, and societal challenges of China's capital, Beijing.

4.2 Task cycle and repetition

Learners engaged in a daily three-step learning cycle: pre-tasks for pre-class preview, core-tasks for in-class learning, and post-tasks for after-class review. Each two-week theme culminated in a summative assessment, namely, a "task in action" in one or a combination of three communicative modes. It concluded with a community-based experiential learning (CBEL) online component that involved live interaction in the target language between learners and native speakers who were graduate students in a program for teaching Chinese as a foreign language. The CBEL was designed as a real-time, live, immersive experience in which learners had to apply learned language elements and functions to interact with native speakers through spontaneous communication.

Framed in this task-based Chinese language course, the task iteration under study involved two expository essays on a societal issue: northern drifters that have posed challenges to population and economic growth, educational equity, and family separation in China. There was a three-week interval between the first and second essay, and the

prompts and requirements for the two essays were identical; both were typed and had to be completed within 20 minutes.

4.3 Participants

Participants included 22 undergraduate students enrolled in two sessions of the same course. Nine participants were heritage learners and 13 were non-heritage learners; 10 were male and 12 were female. The mean age was approximately 21 years old. All took the course purely online.

4.4 Peer review process

During the three-week interval between essays, students went through four steps of peer review. In Step 1, students received the first round of training to get to know the time line, grading criteria and guidelines, and a sample checklist with sample writings, comments, and notes. Step 2 involved a mock peer review practice in which students used the checklist to edit and review another sample essay and discussed their comments and notes. In Step 3, students completed CAPR in groups of three, asynchronously online. Each student reviewed two essays typed by his or her peers and shared the completed checklist for each in a designated Google folder. Each essay was reviewed by two peers in the same group. In Step 4, the peer review activity was conducted in the FFPR format. The same group of students convened to discuss the uploaded checklists and exchange thoughts and feedback based on the checklists. The flow chart in Figure 1 shows the four-step procedure.

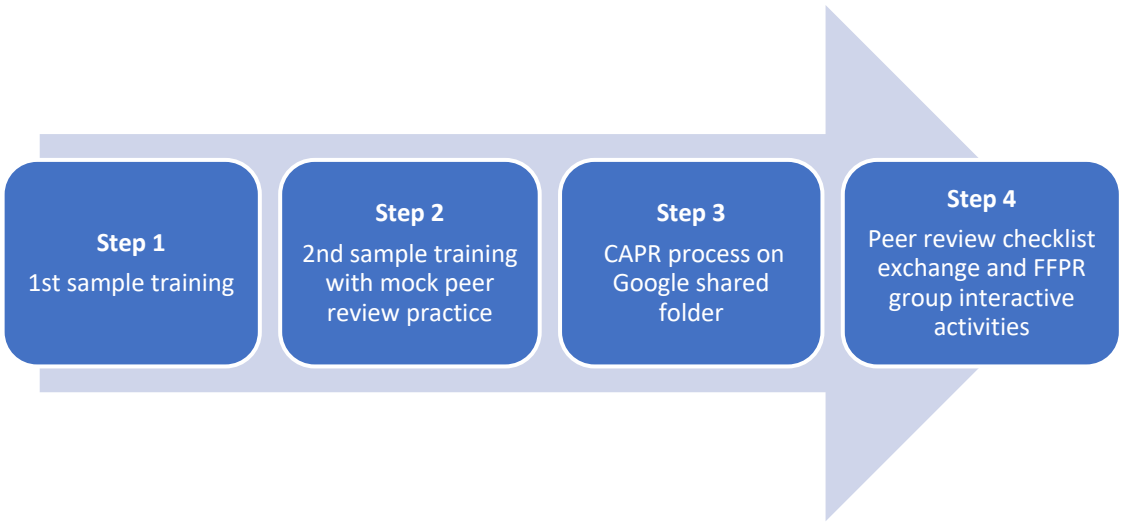

Fig. 1. The four-step procedure of peer review.

5. Data collection and analysis

To answer the first three questions, participants' first and second versions of essays were analyzed and compared in complexity, accuracy, and fluency. Subcategories were defined; based on these, researchers assessed whether writing in the expository essays significantly improved, using quantitative descriptive analysis and statistical analysis through t-tests.

First, complexity in writing was investigated in three dimensions: lexical complexity, syntactical complexity, and discourse complexity. Lexical complexity refers to the number of appropriately used vocabulary words that are required at the advanced level (e.g.: 申办, "to apply," 棘手, "difficult to handle," 迫于无奈, "be compelled against one's will," 苦不堪言, "suffer unspeakably," and 逐年递增, "increase progressively"). Syntactical complexity is denoted by the number of complex and compound sentences in participants' essays.

Syntactical complexity examples:

虽然北京市的就业机会很多，但是生活开销不低。（Complex sentence）

"Although Beijing has a lot of job opportunities, the living expense is high."

天津的学校不错，而且不难申请。（Compound sentence）

"School in Tianjin is good, and applying for it is easy."

Discourse complexity refers to the number of discourse devices, such as organizational devices, coherent devices, and cohesive devices (e.g.: 首先, “first of all,” 其次, “secondly,” 最后, “lastly,” 总而言之, “in conclusion,” and 比如说, “for example”).

Accuracy in writing was analyzed in the subcategories of character accuracy, vocabulary accuracy, and clause accuracy. Character accuracy is defined by the number of correctly typed characters divided by the total number of characters. When calculating character accuracy, multiple instances of the same error were counted only once in one essay. Vocabulary accuracy is defined by the number of correctly used vocabulary words divided by the total word count. Similar to character accuracy, multiple instances of the same vocabulary error were counted only once. Clause accuracy is calculated by the number of correct clauses divided by the total number of clauses. A correct clause was defined as having zero syntactical or grammatical errors.

Lastly, fluency in writing was examined in three subcategories: character count, clause count, and mean length of clause. Character count is the total number of characters excluding punctuation. Clause count is the total number of independent clauses, each containing a semantic unit and with grammatical organization secondary to the sentence. For example, in the sentence below, a total of two clauses were counted:

Clause count example:

他们买不到自己的房子 (clause 1)，户籍也不容易申办 (clause 2)。

“They cannot purchase their own house, and household register is difficult to apply.”

The mean clause length refers to the number of characters in one clause, divided by the total number of clauses.

To answer the fourth research question, an end-of-semester survey was administered to collect learners’ input and comments. The survey included 7-point Likert-scaled questions and an open-ended question that elicited students’ overall comments.

5.1 Did task repetition with peer review as corrective feedback help to improve complexity in L2 writing in Mandarin Chinese? If yes, to what extent?

Results reveal that L2 writing in the repeated task improved in lexical, syntactical, and discourse complexity. The descriptive analysis indicating noticeable change between the first and second versions of the three categories is shown in Table 1.

Table 1. Descriptive data of complexity of the first and second version of essay.

	1st Version				2nd Version			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Lexical	12.10	5.34	5	24	13.64	6.10	5	27
Syntactic	7.28	2.21	4	12	8.77	2.09	6	14
Discourse	1.36	1.5	0	6	1.41	1.26	0	4

T-tests were used to determine significance of the differences between the two essays in the three subcategories of writing complexity. Results are shown in Table 2.

Table 2. Complexity of the first and second versions of essay.

	Lexical complexity	Syntactical complexity	Discourse complexity
Mean of paired differences	-1.545454545	-1.5	-0.045454545
Variance of paired differences	2.354978355	4.166666667	0.997835498
Degree of freedom	21	21	21
t-value	-4.723610907	-3.446737588	-0.213431828
p-value (two-tail)	*0.000115399	*0.002417395	0.83304864

Lexical complexity. As Table 1 indicates, the mean of the appropriate use of required vocabulary at the advanced level in the first and the second versions increased from 12.10 (AD = 5.34) to 13.64 (SD = 6.10). While the minimal use of vocabulary was unchanged in the two versions (min. = 5), the maximal count of vocabulary increased from 24 to 27. T-test results also indicate that lexical complexity significantly increased from the first to the second versions of the essay, $t(21) = -4.72$, $p = .000$. Taken together, these results indicate that through the peer review process, participants were able to use more advanced-level vocabulary.

Syntactical complexity. Analysis of syntactical complexity shows that the mean number of complex and compound sentences increased from 7.28 (SD = 2.21) in the first essay to 8.77 (SD = 2.09) in the second. The minimal use of complex and compound sentences increased from 4 to 6 and the maximal use of complex and compound sentences increased from 12 to 14. The statistical analysis with t-tests also shows that syntactical complexity significantly increased in the second version of essay, $t(21) = -3.44$, $p = .002$. It is therefore justifiable to conclude that through repeated tasks and peer review, participants were using more complex sentences and compound sentences.

Discourse complexity. As far as discourse complexity is concerned, the mean number of discourse devices is 1.36 (SD = 1.5) in the first version of essay and 1.41 (SD = 1.26) in the second. The minimal use of discourse devices remained zero in the second version and the maximal use of the discourse devices slightly decreased from 6 to 4. T-tests show no significant difference in the two versions of essay, $t(21) = -0.21$, $p = .833$. The improvement of discourse complexity in the repeated task is not as significant as the improvement in lexical and syntactical complexity. This may be due to the extra effort needed to reconceive the essay after peer review. The fact that the writing time limit (20 minutes) remained the same in the second essay may have left insufficient time for learners to reorganize their thoughts at the discourse level in the repeated task. The participants were required to write only 300 characters, which was relatively short. To finish the required task, they did not need to use a wide variety of discourse devices to connect, transit, or compare. More importantly, the organizational devices for sequencing ideas that participants had learned

at the advanced-low level under study are not as diversified as those for a higher level of proficiency.

5.2 Did task repetition with peer review as corrective feedback help to improve accuracy in L2 writing in Mandarin Chinese? If yes, to what extent?

Table 3 reveals the accuracy of character, vocabulary, and clause in the first and second version of essay.

Table 3. Descriptive data of accuracy of the first and second version of essay.

	1st Version				2nd Version			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Character	99.24%	0.0055	98.17%	100%	100%	0.0044	98%	100%
Vocab	99.48%	0.0070	97.86%	100%	100%	0.0053	98%	100%
Clause	90.89%	0.0617	75.86%	100%	94%	0.0515	81%	100%

T-tests were conducted to investigate the significance of accuracy improvement from the first version of the essays to the second. Results are shown in Table 4.

Table 4. T-test results of complexity of the first and second version of essay.

	Character Accuracy	Vocabulary Accuracy	Clause Accuracy
Mean of paired differences	-0.004081818	-0.000881818	4.320622727
Variance of paired differences	3.14073E-05	1.04425E-05	416.5508849
Degree of freedom	21	21	21
t-value	-3.416251877	-1.279935022	0.992941484
p-value (two-tail)	*0.002597023	0.214527475	0.332042628

Character Accuracy. The mean accuracy rate of characters increased from 99.24% (SD = 0.0055) in the first essay to 100% (SD = 0.0044) in the second. The minimal accuracy rate decreased slightly from 98.17% to 98%, and the maximal accuracy rate among participants remained 100%. Results from the statistical analysis also indicate that character accuracy in the second version of the essay significantly increased from the first version, $t(21) = -3.41$, $p = .003$. Overall, results show that participants were able to correct their character errors in the repeated task after the peer review process.

Vocabulary Accuracy. The mean accuracy rate of vocabulary increased from 99.48% (SD = 0.0070) to 100% (SD = 0.0053). The minimal accuracy rate increased slightly from 97.86% to 98%, and the maximal accuracy rate among participants remained 100%. The results of the t-test show no significant change from the first to the second version of essay, $t(21) = -1.28$, $p = .215$. The insignificance of this result may be because, among the 11 participants whose vocabulary accuracy rate did not change, 10 had a 100% accuracy rate already. Therefore, results show that some of the participants who had made vocabulary errors in the first essay were able to correct their errors in the repeated task after the peer review process.

Clause Accuracy. Descriptive statistics show positive changes in participants' clause accuracy rate in the second version of essay. The mean accuracy rate of the use of clauses increased from 90.89% (SD = 0.0617) to 94% (SD = 0.0515). The minimal accuracy rate increased from 75.86% to 81%, and the maximal accuracy rate among participants remained 100%. Although descriptive analysis indicates some change, the t-test shows no significant difference in the two versions of essay, $t(21) = 0.99$, $p = .332$. Reasonably, correcting typos in Chinese characters can be readily done by learners themselves through peer interaction. Comparatively speaking, accurate vocabulary use at the lexical level concerns the accurate use of coherent and cohesive devices, idiomatic expressions in context, and a mix of formal and informal expressions, to name a few, and therefore requires higher-level writing skills than lexical items. The production of a clause is the most challenging among the three categories under analysis. A clause that goes beyond the smallest unit of lexical use involves linguistic nuances and a longer length of production.

This further explains the interconnection between complexity and accuracy in the writing process. Correction at the clausal level may not be easily achieved by learners whose proficiency is not yet solid at the advanced level. Attainable growth may rely on accurate input and comments from Chinese speakers who can help clarify the student's original intent and finalize accordingly. The input that student participants received though peer interaction may not have provided sufficient corrective feedback because the student peers were not solid advanced speakers.

5.3 Did task repetition with peer review as corrective feedback help to improve fluency in L2 writing in Mandarin Chinese? If yes, to what extent?

Table 5 reveals descriptive fluency data with respect to character counts, clause counts, and mean clause length.

Table 5. Descriptive fluency data of the first and second versions of essay.

	1st Version				2nd Version			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Character Counts	369.27	79.48	276	574	443.36	82.04	287	581
Clause Counts	30.91	5.71	22	42	36.23	6.00	26	47
Mean Clause Length	11.94	1.50	9.22	16.09	12.25	1.46	8.92	15.7

T-tests were conducted to investigate significant differences in fluency between the two versions of the essay. Results are shown in Table 6.

Table 6. T-test results of fluency of the first and second version of essay.

	Character count	Clause count	Mean clause length
Mean of paired differences	-74.09090909	-5.318181818	-0.308181818
Variance of paired differences	5761.705628	38.41774892	0.213101299
Degree of freedom	21	21	21
t-value	-4.578262934	-4.024468796	-3.131304522
p-value (two-tail)	*0.000163129	*0.000612818	*0.005045259

Character Count. Results show that the mean character count per essay increased from 369.27 (SD = 79.48) to 443.36 (SD = 82.04). The minimal character count increased from 276 to 287, and the maximal character count increased from 574 to 581. Statistical analysis also indicates that fluency in character count significantly increased in the second version of writing, t(21) = -4.57, p = .000. Within the required 20-minute limit, participants were able to write more characters, which shows observable improvement in writing fluency.

Clause Count: As the table shows, the clause count per essay increased from 30.91 (SD = 5.71) to 36.23 (SD = 6.00). The minimal clause count increased from 22 to 26, and the maximal clause count increased from 42 to 47. The statistical analysis shows that fluency in terms of clause count significantly increased in the second essay, t(21) = -4.02, p = .001. Participants tended to compose more independent clauses in the second essay within the limited time, indicating increasing fluency in clauses in writing.

Mean Clause Length. The mean clause length per essay increased from 11.94 (SD = 1.50) to 12.25 (SD = 1.46) per essay. The minimal mean clause length decreased from 9.22 to 8.92, and the maximal mean clause length decreased from 16.09 to 15.7. T-test results indicate that fluency in terms of mean clause length significantly increased in the repeated writing task, t(21) = -3.13, p = .005. The examples below show how the mean clause length increased in the second version of the essay:

Longer clause length, example (1)

1st version: 虽然北京经济不错，人口稠密，所以找机会比较难。

“Although Beijing’s economy is good, the population is large, so it is difficult to find opportunities.”

2nd version: 虽然北京经济不错，但是人口稠密，所以找机会比较难。

“Although Beijing’s economy is good, but* the population is large, so it is difficult to find opportunities.”

Longer clause length, example (2)

1st version: 去天津上学不好，因为孩子早上得自己坐火车半个小时上学，晚上坐火车半个小时回家。

“Attending school in Tianjin is not good as children have to take the half-hour train in the morning to go to school and at night to go back home.”

2nd version: 去天津上学不会消除北京学生们不平等的问题，因为孩子得早上自己坐火车半个小时上学，晚上坐火车半个小时回家。

“Attending school in Tianjin will not eliminate the inequality among Beijing children as they have to take the half-hour train in the morning to go to school and at night to go back home.”

In the first example, the participant added 但是 (but) in the second version of essay; this word is necessary in the structure of Mandarin Chinese. In the second example, the participant further elaborated the idea in the sentence to enrich its meaning. In the second iteration of the essay, the majority of learners deliberately enriched and expanded their thoughts.

Some learners produced slightly shorter clauses in the second versions of their essays; the following three examples investigate why. The minor increase of the mean clause length may result from grammatical corrections and formal choice of vocabulary that the participants made after peer review. For example, after participants corrected vocabulary or grammar, the length of their sentences sometimes decreased by one to three characters, as exemplified below.

Shorter clause length, example (1)

1st version: 搬到别的地方也很苦，所以北漂族搞得人心惶惶。*

“Moving to other areas is also difficult so the northern drifters made* anxious.”

2nd version: 搬到别的地方也很苦，所以北漂族会人心惶惶。

“Moving to other areas is also difficult so the northern drifters are anxious.”

Shorter clause length, example (2)

1st version: 他们的生活会有比较少的压力。

“There will be less pressure in their life.”

2nd version: 他们的生活压力会降低。

“Their life pressure will be reduced.”

In the first example, the participant corrected the use of 搞得 (make) and switched to 会 in the second version, decreasing the length of the clause by one character. In the second example, the participant used an advanced-level word, 压力会降低 (pressure will reduce), to paraphrase the predicate in the previous sentence, 有比较少的压力 (has less pressure), which decreased the character count in the second sentence. Another reason the mean clause length declined in some participants’ second essays is that they corrected their use of punctuation, which often reduced the length of each clause, as shown in the following third example.

Shorter clause length, example (3)

1st version: 尽管如此，大家还是想来这些城市因为大家都希望能安居乐业。*

“Despite all this, people still want to come to these cities because they want to have a nice job and stable life.”

2nd version: 尽管如此，大家还是想来这些城市，因为大家都希望能安居乐业。

“Despite all this, people still want to come to these cities, because they want to have a nice job and stable life.”

5.4 What were learners' input and comments on the peer review process as a means of intervention between the two tasks?

The results from the survey, both the 7-point Likert-scaled questions and the open-ended question, suggest that most learners had positive perceptions of the peer review process. Typical comments included observations that having "peer reviews sooner in Chinese curriculum would have been so beneficial" and "getting feedback from classmates was really encouraging and helpful." Benefits of peer review cited by participants include the enhancement of peer interaction, writing improvement quality and motivation, specific benefits of CAPR and FFPR, the effectiveness of the peer review training process, and minimizing pressure.

As shown in Table 7, participants believed that the peer review process significantly enhanced the amount and quality of peer interaction. Most of the participants agreed (and 36.36% strongly agreed) that the peer review process promoted peer learning, and most also agreed (and 45.45% strongly agreed) that it enhanced peer interaction. In the open-ended questions, participants' responses included "peer review is a great way to enter engaging discussion around Chinese writing," and it was "an interesting and inclusive way to engage [not only] with the learning material but also with peers." Some participants mentioned that the peer review process was very helpful especially in online classes, as "interacting with classmates was hard to achieve in online format" compared with face-to-face classes.

Table 7. Participants' perspectives about peer review (PR)'s effect on peer learning and interaction.

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
PR promoted peer learning.	0 (0.00%)	1 (4.55%)	0 (0.00%)	0 (0.00%)	6 (27.27%)	7 (31.82%)	8 (36.36%)
PR enhanced peer interaction.	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	6 (27.27%)	6 (27.27%)	10 (45.45%)

Most participants also believed that the peer review process helped and motivated them to improve their writing, as indicated in Table 8. Generally regarding the benefits in writing preparation and improvement, many of them thought (and 36.4% strongly believed) that the peer review process prepared them to write better. Most also believed (and 40.9% strongly believed) that the peer review process motivated them to better their writing. They also expressed (and 59.1% strongly agreed) that they appreciated the opportunity to get feedback from peers through the process.

Table 8. Participants' perspectives about PR in writing improvement and motivation.

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
PR prepared me to write better.	0 (0.00%)	0 (0.00%)	2 (9.09%)	3 (13.64%)	2 (9.09%)	7 (31.82%)	8 (36.36%)
PR motivated me to better my writing.	0 (0.00%)	0 (0.00%)	0 (0.00%)	2 (9.09%)	6 (27.27%)	5 (22.73%)	9 (40.91%)
Appreciation of peer feedback.	0 (0.00%)	0 (0.00%)	0 (0.00%)	2 (9.09%)	1 (4.55%)	6 (27.27%)	13 (59.09%)

In the open-ended question, participants mentioned that peer review was very helpful, in that it was "really good to have others notice things about my writing that I had not realized." They also reported that peer review was "incredibly helpful" for them in "providing fresh perspectives in writing." One participant wrote,

> *We often fall back on the same vocabulary and grammatical structures in our writing, but the insight of two fresh pairs of eyes allows to me the opportunity to view their thought process and ultimately create richer writing.*

Some participants also mentioned that the peer review process helped to improve not only their writing but also their reading skills, as they needed to "identify areas for improvement in my peers' essays" in order to provide comments and suggestions. They also commented that their vocabulary and grammar skills improved during that process. For example, one participant expressed a reflection about vocabulary and grammar in the process of peer review:

Getting feedback from my peers has been enriching because it is interesting to see the unique writing style of each of my peers. They all speak differently and use different grammar and vocabulary to speak about similar topics, and it has given me inspiration on different ways I can expand my vocabulary and grammar usage.

Regarding the effectiveness of having both asynchronous checklists (CAPR) and face-to-face discussion (FFPR) in their peer review process, participants held positive perceptions: 22.73% of them very much agreed and 50.00% strongly agreed that getting feedback from the checklists completed by peers was helpful for improving their writing, according to Table 9. In addition, 27.27% very agreed and 31.82% strongly agreed that discussing with peers at the follow-up face-to-face virtual meeting was helpful for improving their writing.

Table 9. Participants' perspectives about the use of asynchronous checklist (CAPR) and face-to-face discussion (FFPR).

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
Effectiveness of completing the checklist.	0 (0.00%)	0 (0.00%)	1 (4.55%)	1 (4.55%)	4 (18.18%)	5 (22.73%)	11 (50.00%)
Effectiveness of face-to-face session with peers.	1 (4.55%)	0 (0.00%)	0 (0.00%)	2 (9.09%)	6 (27.27%)	6 (27.27%)	7 (31.82%)

Participants also reported that the peer review process helped to minimize the pressure they felt around writing, as shown in Table 10. Among 22 participants, 31.82% very agreed and 36.36% strongly agreed that the peer review process minimized their pressure. In responses to the open-ended question, a typical comment was that they found peer review to be "a great way to revise and improve my writing skills in a low-pressure way." Some

also mentioned that they preferred having more than one partner in the peer review group, as it further reduced their hesitancy and pressure in giving feedback to others. One wrote,

> *I was a little hesitant to give advice sometimes if I wasn't sure if my advice was correct or not but having a second classmate really helped.*

Table 10. Participants' perspectives about peer review's effect on pressure reduction.

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
PR's effect on minimizing pressure.	0 (0.00%)	1 (4.55%)	0 (0.00%)	0 (0.00%)	6 (27.27%)	7 (31.82%)	8 (36.36%)

Participants also rated the peer review training process. Among 22 participants, 18.18% very agreed and 54.55% strongly agreed that the peer review template and examples were helpful for their preparation in reviewing their peer's essay. In addition, 31.82% of the participants very agreed and 40.91% strongly agreed that the "mock" peer review practice was helpful in preparing them to complete the real checklist.

Table 11. Participants' perspectives on training for the peer review process.

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
Effectiveness of PR template and examples.	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (4.55%)	5 (22.73%)	4 (18.18%)	12 (54.55%)
Effectiveness of "mock" PR.	1 (4.55%)	0 (0.00%)	0 (0.00%)	2 (9.09%)	3 (13.64%)	7 (31.82%)	9 (40.91%)

In conclusion, regarding the first three research questions, the repeated writing task with peer review improved students' writing performance in all three aspects: complexity, accuracy, and fluency. However, the extent of improvement varies in the three aspects. For participants' complexity improvement, their lexical and syntactical complexity improved significantly. However, no obvious change in participants' discourse complexity was found. In terms of writing accuracy, participants' character accuracy significantly improved. Comparatively speaking, participants' vocabulary and clause accuracy rates did not change significantly. This is partially due to the fact that half of the participants had already reached 100% accuracy in vocabulary use in the first version, and they maintained the highest accuracy rate in the second version. Similarly, the change in clause accuracy is not significant. Despite some progress in accurate reconstruction of clauses, the degree of improvement does not indicate significant growth in the repeated task. With respect to fluency, participants were able to elaborate further based on their peers' comments. This contributes to the increased character count, clause count, and mean clause length in the second versions. While quantity improved, there is room for further improvement in quality in terms of the accurate use of formulatic and idiomatic expressions and grammatical structures, as participants progress toward the advanced level.

The survey results provide valuable insight into learners' perspecitves on the effect of corrective feedback through peer review. Positive results were also found in the answers to the fourth research question. Most participants clearly held positive perceptions of the design and implementation of the peer review process. Specifically, they appreciated the pedagogical value of peer review in the design and implementation process, frequent interaction, positive effect on their motivation, pressure reduction, and the combination of asynchronous and synchronous discussion. Results also show that the preparatory training and self-regulation contributed to an effective peer review process.

6. Pedagogical Implications

Task repetition enhances learners' performance. The remedial or interventive strategy used for the repeated task should be well-thought-out in order to maximize pedagogical

value. Corrective feedback through peer review is a socio-cognitive approach to writing that yields benefit for L2 writing, enabling learners to engage with, reflect on, and interpret written texts through interactional exchanges about the task repetition under study. In whatever format, it is worth considering the incorporation of this practice in the curriculum for L2 writing to foster active and student-centered learning and to diversify ways of essay grading and commenting. Maximizing the effects of this type of peer learning and interaction requires careful planning, preparation, and implementation. In light of the design of the peer review process in this study, the following strategies are highly recommended.

First, conduct training ahead of time. Training and preparation are crucial to ensure that learners know what and how to comment on their peers' writing, and, most importantly, to understand teachers' expectation of their performance. This requires the instructor to plan ahead, in order to give students very clear step-by-step instructions and impart readiness skills. Students' familiarity with expectations and knowing how their work is evaluated will help relieve their tension and remove unnecessary psychological burdens. Instructors should consider giving a completion grade, a letter grade, or numerical points in consultation with students. Note that one round of practice is insufficient. It is recommended that the instructor invite learners to practice analyzing writing samples produced by themselves or their peers in addition to teacher-generated samples annotated by the instructor. Working on peers' work also fosters peer bonding in the learning process.

Second, combine asynchronous with synchronous interaction. A combination of asynchronous and synchronous feedback is highly recommended. It offsets the disadvantages of each format alone to achieve the best outcomes. A checklist is pivotal: providing samples, annotations, and explanatory notes helps learners get a good start in preparation for the peer review process. The process of completing a checklist fosters deliberate thinking, as it gives learners time to consider and search for language resources they need. This self-paced option strategically caters to individual differences, accommodates individual needs for self-learning, and potentially promotes more turn taking, meaningful discussion, and peer interaction. Following it with face-to-face discussions yields great value, as students can get points and comments clarified, justified,

and confirmed whenever needed. This also creates an opportunity for students to think further about decisions to maintain their original writing or modify it.

Third, create proximal group dynamics. Consider forming a group of three rather than a two or four. Grouping three members together allows each student to comment on two entries and receive advice and feedback from two peers. While forming a group of four helps with the richness of feedback, it increases students' workload and the length of meeting time and is therefore not strongly encouraged. As in student-centered communicative tasks, the importance of grouping strategies for peer review cannot be understated. Avoid forming high-proficiency and low-proficiency groups: it is more effective to have a mix of proficiency levels in one group, unless differentiated instruction is needed for a multi-leveled class. When grouping students, socio-cognitive factors are worthy of full consideration, including but not limited to personality, gender, behavior, cultural background, communication patterns, and so on. As a final note, it is a good practice to create learning space and social space at the same time. Although learners' interactional feedback exchanges add vital momentum in the learning space, interactional feedback exchanges in the social space matter as well. Encouraging learners to express appreciation, welcoming, praise, and social chatting contributes to group dynamics and builds strong support for productive and proximal textual exchanges in whatever form.

7. Conclusion

This study offers preliminary results pointing to the positive effects of task repetition mediated with peer review as a means of corrective feedback. Learners' writing improves in complexity, accuracy, and fluency, to various degrees. In terms of fluency, the expository essay under study significantly increases in length and exhibits more elaboration and supporting ideas. Complexity increases at lexical and syntactic levels, indicative of increasing use of complex and compound sentences, but not at the discourse level, meaning the reorganization of paragraphing and connecting of thoughts. Accuracy has been much less frequently studied in research on L2 writing. This study fills that gap and reveals that the accuracy rate increases in typed character production, but not in vocabulary or clause.

These preliminary results need to be corroborated through further studies with refined experimental designs that include both control and experimental groups. Furthermore, groups featuring no feedback versus different types of feedback, such as instructor's feedback, need to be examined.

This study involves only expository essays; thus, there is also a need for investigations that focus on other genres of writing such as descriptive, narrative, and argumentative, and at different levels of proficiency. Character input remains an area that needs further investigation. The distinction between handwritten and typed essays in relation to performance in L2 writing deserves extensive research to expand our understanding of how these two types of input link to different aspects of performance in L2 writing; this calls for studies that compare the results of typed essays with those of handwritten essays through different research designs. Turning from complexity, accuracy, and fluency to the effects of peer review on L2 writing, future studies should also examine the extent to which reviewers' suggestions impact peers' revision. It is worth analyzing the types of peer review comments and revisions made after peer review, and the factors affecting the incorporation of peer review comments.

Repeating a writing task with peer review, if carefully designed and guided, leads to improvement in L2 writing. Coupled with the strengths of pedagogical advancement in task repetition, this is an emerging area that invites more research in teaching Chinese as a foreign language.

Acknowledgement

The study was made possible by funding from the National Endowment for the Humanities awarded to Miao-fen Tseng for her Daniels Family NEH Distinguished Teaching professorship. Special thanks to Yujia Sun for her analysis of language data based on typed essays. Credits should also be given to students who participated in the study in spring 2021 in the 3-year initiative on the creation of a technology-mediated task-based language course.

References

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge, UK: Oxford University Press.

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*, 2-20. Http://dx.doi.org/10.1016/j.jeap.2012.01.001.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*, 5-35. Http://dx.doi.org/10.5054/tq.2011.244483.

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond T-units-A response to WeiWei Yang. *TESOL Quarterly, 47*, 192-201. Http://dx.doi.org/10.1002/tesq.84.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willils & D. Willis (Eds.) *Challenges and change in language teaching*, 136-146.

Bygate, M. (1999). Task as context for the framing, reframing, and unframing of language. *System, 27*(1), 33-48.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M Swain (Eds.) *Researching pedagogic tasks: Second language learning, teaching and testing,* pp. 23-48. Harlow: Pearson Education.

Bygate, M. (2018). *Learning language through task repetition.* Philadelphia, PA: John Benjamins.

Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language*, pp. 37-74. Philadelphia, PA: John Benjamins.

De Guerrero, M. C. M., & Vilamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal, 84*, 51-68.

de, John, N., & Perfetto, C. A. (2011). Fluency training in the ESL classroom: an experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533-568.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford, UK: Oxford University Press.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *A Journal of Research in Language Studies, 54*(2), 227-275.

Ellis, R. (2008). *The study of second language acquisition.* Oxford: Oxford University Press.

Fotos, S., & Hossein, N. (Eds.) (2007). *Form Focused Instruction and Teacher Education: Studies in Honour of Rod Ellis*. Oxford: Oxford University Press.

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review, 61*(3), 325-353.

Hammerly, H. (1991). Fluency and accuracy toward balance in language teaching and learning. *Multilingual Matters*.

Hayes, J., R., & Flower, L., S. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg, & E. R. Steinberg (Eds.) *Cognitive process in writing*, pp. 31-50. Hillsdale: Lawrence Erlbaum Associates.

Hanjani, A. M., & Li, L. (2014). Exploring L2 writers' collaborative revision interactions and their writing performance. *System, 44*, 101-114.

Ho, M. C., & Savignon, S. J. (2007). Face-to-Face and Computer-Mediated Peer Review in EFL Writing. *CALICO Journal, 24*(2), 269-290. Retrieved March 29, 2022 from https://www.learntechlib.org/p/74268/.

Hossein, N., & Fotos, S. (2011). *Teaching Grammar in Second Language Classrooms: Integrating Form-Focused Instruction in Communicative Context*. New York, NY: Routledge.

Hu, G. W. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research, 9*(3), 321-342.

Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and Issues.* Cambridge: Cambridge University Press.

James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. London: Routledge.

Johnson, M. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing, 37,* 13-38.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens. A. (2013). Working memory in written composition: An evaluation fo the 1996 model. *Journal of Writing Research, 5*, 159-190. Http://dx.doi.org/10.17239/jowr-2013.05.02.1

Kern, R. (2000). *Literacy and language teaching*. Oxford: Oxford Applied Linguistics.

Kim, Y., Crossley, S., Jung, Y., Kyle, K., & Kang, S. (2018). The effects of task repetition and task complexity on L2 lexicon use. In M. Bygate (Ed.), *Learning language through task repetition*. Philadelphia, PA: John Benjamins.

Kim, Y., & Ventura-Tracy, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based instruction? *System, 41*, 829-841.

Khezrlou, S. (2020). The role of task repetition with direct written corrective feedback in L2 writing complexity, accuracy and fluency. *Journal of Second Language Studies, 3*(1), 31-54. DOI:10.1075/jsls.19025.khe

Larsen-Freeman, D. (2018). Task repetition or task iteration? In M. Bygate (Ed.), *Language learning through task repetition*, pp. 311-329. John Benjamins.

Levelt, W. J. M. (1989). *Speaking, from intention to articulation*. Cambridge, MA: The MIT Press.

Levi Altstaedter, L. (2016). Investigating the impact of peer feedback in foreign language writing. *Innovation in Language Learning and Teaching*, 1-15.

Long, M. H. (1988). 'Instructed interlanguage development'. In L. Beebe (ed.), *Issues in Second Language Acquisition: Multiple Perspectives*, 115-141. Rowley, MA: Newbury House.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching, 40*(3), 243-249.

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Special issue Tasks in language pedagogy. Language Teaching Research, 4*(3), 221-250.

Lynch, T., & Maclean, J. (2001). A case of exercising: Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.),

Researching pedagogic tasks; Second language teaching and learning, pp. 141-162. Harlow: Pearson Education.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*(1), 27-56.

McGroarty, M. E., & Zhu, W. (1997). Triangulation in classroom research: A study of peer revision. *Language learning, 47*(1), 1-43.

Min, H. T. (2005). Training students to become successful peer reviewers. *System, 33*(2), 293-308.

Nita, R., & Baba, K. (2014). Task repetition and L2 writing development. In H. Brynes & R. Manchón (Eds.), *Task-based Language Learning Insights from and Form L2 Writing,* pp. 107-136.

Pica, T., Kanagy, R., & Falodun, J. (1993). 'Choosing and using communication tasks for second language instruction. In S. Gass and G. Crookes (Ed.), *Task-Based Learning in a Second Language*, 9-34. Clevedon: Multilingual Matters.

Plough, I., & Gass, S. (1993). Interlocutor and task familiarity: Effects on interactional structure. In G. Crookes & S. Gass (Eds.), *Task and language learning: integrating theory and practice*, pp. 35-56. Clevedon: Miltilingual Matters.

Revesz, A., Sachs, R., & Hama, M. (2012). *Task Complexity, Type-Token Frequency, and Individual Differences on the Acquisition of Past Counterfactuals*. Unpublished paper presented at American Association for Applied Linguistics 2012 Conference, Boston, MA.

Richards, J. C. (2001). *Curriculum Development in Language Teaching*. Cambridge, UK: Cambridge University Press.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27-57. Http://dx.doi/10.1093/applin/22.1.27.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL-International Review of Applied Linguistics in Language Teaching, 43,* 1-33. Http://dx.doi.org/10.1515/ital.2005.43.1.1.

Robinson, P. (2011). *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. Amsterdam: John Benjamins.

Saeed, M., Ghazali, K., & Aljaberi, M. (2018). A review of previous studies on ESL/EFL learners' interactional feedback exchanges in face-to-face and computer-assisted peer review of writing. *International Journal of Educational Technology in Higher Education, 15*(6). DOI:10.1186/s41239-017-0084-8.

Samuda, V., & Bygate, M. (2008). *Tasks in Second Language Learning*. Basingstoke, UK: Palgrave MacMillan.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University Press.

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction*, pp.183-205. Cambridge, UK: Cambridge University Press.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82,* 320-337.

Swain, M. & Lapkin, S. (2002). Talking it through: two French immersion learners' response to reformulation. *International Journal of Educational Research, 37*(3), 285-304.

Tseng, M. (2014). *The Handbook of Tasks and Rubrics for Teaching Mandarin Chinese* (Volume I), Chicago, IL: The Phoenix Tree.

Tseng, M. (2017). Chinese language teachers' perceptions of tasks and task-based language teaching. *Journal of National Council of Less Commonly Taught Languages,* 29-65.

Tseng, M. (2018). *The Handbook of Tasks and Rubrics for Teaching Mandarin Chinese* (Volume I, E-version added), Chicago, IL: The Phoenix Tree.

Tseng, M. (2019). Creating a task-based language course in Mandarin Chinese. In Chris Shei, Monica Zikpi, and Der-Lin Chao (Eds.), *The Routledge Handbook of Chinese Language Teaching*, pp. 118-133.

Tseng, M. (2020). *The Handbook of Tasks and Rubrics for Teaching Mandarin Chinese* (Volume II), Chicago, IL: The Phoenix Tree.

Tseng, M. (2020). *The Handbook of Tasks and Rubrics for Teaching Mandarin Chinese* (Volume III), Chicago, IL: The Phoenix Tree.

Van den Brandon, K., Bygate, M., & Norris, J. (Eds.) (2009). *Task-Based Language Teaching: A Reader.* John Benjamins.

Wang, Z., & Chen, G. (2018). Discourse performance in L2 task repetition. In M., Bygate (Ed.), *Learning language through task repetition*. Philadelphia, PA: John Benjamins.

Willis, D. (1993). Comments on Michael H. Long and Graham Crookes's 'Three approaches to task-based syllabus design'. *TESOL Quarterly, 27*(4), 726-733.

Willis, J. (1996). *A Framework of Task-Based Language Teaching*. Harlow: Longman.

Willis, D., & Willis, J. (2007). *Doing Task-Based Teaching: A Practical Guide to Task-Based Teaching for ELT Training Courses and Practicing Teachers*. Oxford, UK: Oxford University Press.

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

Yang, Y. F. (2011). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology, 42*(4), 687-700.

Zhu, W. (1995). Effects of Training for Peer Response on Students' Comments and Interaction. *Written Communication, 12*(4), 492-528.

〔收稿：2022.04.14 修訂：2022.05.03〕

曾妙芬 Miao-fen Tseng
美國維吉尼亞大學東亞語言文化與文學系
Department of East Asian Languages, Literatures, & Cultures
University of Virginia
mt3z@virginia.edu

耿子怡 Ziyi Geng
美國維吉尼亞大學東亞語言文化與文學系
Department of East Asian Languages, Literatures, & Cultures
University of Virginia
zg8md@virginia.edu

華語寫作中的任務重複循環：以生生互動回饋作為糾錯手段之研究報告

曾妙芬*

耿子怡

美國維吉尼亞大學

摘要

任務的重複與循環在口語提升方面的研究已經有了豐碩的成果(Bygate, 2001)。然而,其對於華語作為第二語言寫作的影響仍為一個未開發的領域。本文根據 Bygate（1999）的框架，探討在溝通式任務重複與循環的情境下，學習者語言建構與重塑的情況，以及其對寫作說明文體的影響。在任務重複循環的過程中，以同儕之間生生互動與回饋作為修改寫作的手段。實驗參與者為 22 位選修三年級中文課程的大學生，他們透過同儕之間線下與線上兩者結合的形式完成寫作回饋的互動、交流與修稿，並於寫作任務完成後，填寫一份針對生生寫作回饋設計與教學目標的問卷調查。寫作語料分析包括描述性統計與 t-test 的推論性統計。分析結果顯示，學生全程完成寫作任務後，在複雜度、正確度與流利度三方面都呈現了不同程度的提高。問卷結果亦肯定生生交流回饋作為有效寫作糾錯手段對寫作產生的價值與效能。本文最後提出本研究對寫作任務的啟發，實驗結果的限制，以及未來研究方向。

關鍵詞：任務的重複與循環、生生互動回饋寫作、第二外語寫作、複雜度，正確度，流利度、任務教學

* 本文通訊作者。