RESEARCH   **Open Access**

Effects of input text genre on Chinese young EFL learners' performance on continuation tasks

Xueping Li[1*] and Yiru Wang[1]

*Correspondence:
lixueping01@tyut.edu.cn

[1] College of Foreign Languages, Taiyuan University of Technology, 209 University Street, Yuci District, Jinzhong City, Shanxi Province, China

Abstract

This study examines how the genre of input text influences Chinese young EFL learners' performance on continuation tasks. Participants were 30 students in Grade 9 and repeated-measures design was adopted to compare their performance on narrative and argumentative continuation tasks, in terms of writing quality, CAF indicators as well as alignment. Results showed that (i) genre had no significant influence on students' scores on the two continuation tasks; (ii) argumentative continuation task brought about more complex sentences and words, as reflected by most indicators, but narrative continuation task led to relatively longer writing, and no significant difference was found for accuracy; (iii) students' writings on argumentative task contained more 4-word phrases that were drawn from the input text, while the writings on narrative task were semantically more related to the reading input. This study provides another piece of validity evidence for employing continuation tasks in different genres in language tests and has implications for the writing teaching of Chinese young EFL learners.

Keywords: Genre, Continuation task, Writing quality, CAF, Alignment

Introduction

As an integrated writing task, the continuation task has received a great deal of attention in language teaching and assessment (Shi et al., 2020). It requires learners to read an incomplete text and then write a continuation in a logical and coherent way (Zhang & Wang, 2022). Many scholars believe it to possess great learning potential due to its inherent interactive alignment (Pickering & Garrod, 2004). Besides, since it can effectively integrate language input and output, imitation and creation, as well as learning and application (Wang, 2012), it well suits China's context of English teaching and learning, which is characterized by more reading and writing but less listening and speaking (Jiang & Chen, 2015).

Previous studies have investigated the effects of continuation task on learning from different perspectives; however, most of them were situated in the contexts of tertiary education and high school, and few have explored its pedagogical implications on younger learners (e.g. middle school), who mainly face two challenges in EFL writing: to write correctively and creatively (Copland et al., 2014). Moreover, since continuation tasks are considered to be authentic and capable of reducing content-bias and bringing

about positive washback, they are increasingly adopted in many large-scale, high-stakes language tests (Shi et al., 2020). However, only a few of previous studies on task-related variables have examined the influence of input text genre, which is a key consideration in designing both reading and writing tasks.

Therefore, this study adds to the current research efforts by investigating how EFL young learners (i.e. middle school students) perform on continuation tasks with input texts in different genres, so as to explore its learning potential in the writing teaching of young EFL learners, and on the other hand to provide more validity evidence for adopting the continuation task in various English tests.

Literature review

Genre and ESL/EFL writing

Genre refers to the socially conventionalized ways of language use that can serve specific communicative purposes (Berman, 2008; Grabe, 2002; Swales, 1981). Generally, narratives are agent-oriented, and describe events and actions performed by people according to the temporal sequence (Berman & Nir-Sagiv, 2007); while argumentatives are topic-oriented, and express an opinion on the basis of facts and logic (Grabe, 2002). These differences will influence not only how writers organize the two types of writings, but also the specific language features (Qin & Uccelli, 2016).

Under the CAF (Complexity, Accuracy and Fluency) framework that measures L2 writing development (Wolfe-Quintero et al., 1998), many scholars have examined how genre influences ESL/EFL writing, mostly in terms of syntactic complexity (Zhang & Liu, 2021). A prevailing conclusion has been that argumentatives are syntactically more complex than narratives, though by means of different indices (e.g. Bi, 2020; Lu, 2011; Yoon & Polio, 2016; etc.). In terms of lexical complexity measured by D, Yoon (2021) found narratives were more complex than argumentatives, quite opposite to the results of Qin and Uccelli (2016), while Wang et al. (2022) found no significant differences. As for accuracy and fluency, only Wang et al. (2022) found that argumentatives contained more errors than narratives, and Yoon and Polio (2016) found no significant difference in fluency between Chinese college students' narrative and argumentative compositions.

Results of the effects of genre on writing scores are not consistent either, probably due to factors such as participant type, task type and condition, etc. For example, in contrast to the findings of Qin and Uccelli (2016) and Shao (2021) that genre did not influence students' scores on narrative and argumentative compositions, Ye and Yan (2010) found non-English majors' scores on narratives were significantly lower than those on argumentatives, while Zhang and Liu (2021) found that college students' scores on narratives and argumentatives differed with whether they were required to complete their compositions within the time limit. Using summary writing as the task, Li (2014) found that students performed better in expository text than in narrative texts.

According to the theoretical writing model of Kellogg (1996), there are six processes involved in writing: planning, translating, programming, executing, reading and editing, all requiring the support of ESL/EFL learners' limited working memory. When learners are writing compositions in different genres, in order to fulfill the targeted communicative purposes, they need to plan different organizations and use appropriate languages. However, their "incomplete and unautomatized linguistic system and lack of genre and

discoursal knowledge" become the extra cognitive burden (Li, 2023, p. 648). Therefore, genre is an important factor that may influence ESL/EFL learners' writing performance.

The continuation task

It is believed that the continuation task can enhance student learning because of the induced alignment between the input text and L2 learners in terms of situational models and linguistic representations (Wang & Wang, 2015). Peng et al. (2020) construed alignment "as a complex process in which learners engage in coordinated interaction with the input text" (p. 367). There have been prolific empirical studies indicating the effectiveness of this task. For example, it can generate more gains measured by indices of CAF (Jiang & Chen, 2015), improve lexical accuracy (Wang & Wang, 2015), help with students' acquisition of a certain grammatical structure (e.g. Jiang & Tu, 2016; Sun & Wang, 2018, 2022; Wang & Cao, 2020; Wang & Wang, 2019; Xin, 2017; etc.). Besides, it also has positive influence on students' writing coherence (e.g. Peng, 2017; Zhang, 2016), written rhetoric (Yang, 2018a), writing anxiety and proficiency (Zhang & Qin, 2020).

As one type of integrated writing tasks, the continuation task may influence test-takers' writing performance due to its various task characteristics (Bachman, 1990). In terms of different task conditions, Shi et al. (2020) investigated how four different prompts influenced students' writing scores and strategy use, and Xiao (2013) examined whether the presence of input text during the writing process influenced the accuracy of students' writings. In terms of the characteristics of input text, previous studies have considered factors such as topic familiarity (Gu et al., 2022), linguistic complexity (Peng et al., 2020; Xin & Li, 2020), language (L1 vs. L2) (Wang & Wang, 2014), and students' interest (Xue, 2013).

Despite that genre is essential in L1 and L2 writing research (Tardy, 2006), only a few have touched upon how it may influence students' writing in the continuation task. For example, Guo and Wang (2019) found genre only influenced English majors' scores on structure, but not on content, language and the total score. However, Zhang et al. (2023) found no difference in high school students' scores on narrative, argumentative and expository continuation tasks, quite similar to the results of Zhong (2021). Zhang and Zhang (2017) did not find any difference either in lexical or phrasal alignment in English majors' continuation writings, but Zhang et al. (2023) found that narrative continuation task produced more alignment effects than argumentative one by means of indices of LSS (Latent Semantic Similarity) and LSM (Language Style Matching).

As mentioned earlier, ESL/EFL learners' limited knowledge of language, genre and discourse may demand more processing loads when they are writing compositions of different genres. However, in the case of continuation task that combines reading and writing, the input text may serve as the "scaffold" and hence influence learners' writing process and outcome. Yet, results in this regard are quite inconclusive, both in terms of scores and alignment effects.

Research gaps and research questions

To sum up, as a type of learning task, continuation tasks can enhance student learning at various levels. This effect is estimated to be medium according to the meta-analysis of Ren and Lv (2021), with genre serving as an important moderator. Meanwhile, as a

type of integrated writing task, continuation tasks are increasingly adopted in large-scale English tests, which means that validity evidence about this type of task should be collected on a continual basis. However, previous studies have been mainly conducted in tertiary education and high school and focused on narrative continuation tasks, and more importantly, results of cross-genre comparisons have been inconsistent.

Therefore, this study aims to extend the current research by examining how younger EFL learners from middle school perform on continuation tasks with input texts in different genres, so as to explore its learning potential in the writing teaching of this group of learners, who may differ with learners in high school and university in the following ways.

First is their English proficiency level. According to CSE (China's Standards of English Language Ability, Ministry of Education of the People's Republic of China, 2018), middle school students are roughly at Level 3 of the Elementary Stage, while high school students are at Level 4 of the Intermediate Stage and university students' English proficiency levels may range from Level 5 of the Intermediate Stage to Level 7 of the Advanced Stage (Zhu & Cao, 2020). This means that the linguistic system of this younger group is more limited than that of students in high school and university, which may result in different patterns of using the input text as the "scaffold" (i.e. alignment). The second and related one is their working memory capacity. Generally speaking, one's working memory capability increases across the life span (Cowan, 2014). So it is reasonable to assume that middle students may have more limited working memory, which, in combination with their lower level of English proficiency, will cause greater cognitive burden when they complete the continuation task and give rise to different results from previous studies. The third is their previous experience with continuation task. Since the continuation task has been adopted in NMET (National Matriculation Entrance Test) for several years, high school students are quite familiar with this task type; however, middle school students have done little exercise of this type, since it has not appeared in the senior high school entrance examination of the province in the present study.

Based on these considerations, this study aims to address the following research questions:

(1) How does genre influence Chinese young EFL learners' scores on continuation tasks?
(2) How does genre influence Chinese young EFL learners' continuation in terms of CAF?
(3) Will input texts in different genres produce different alignment effects?

Methods

Participants

Participants included 30 nine-graders (16 males and 14 females) from a middle school in northern China, with an average age of 15 years old. They were from two intact classes and taught by the same English teacher. Two months before this study, the participants had taken an English test that was administered by the local municipal government to nine-graders of several middle schools and aimed to prepare students for the coming

provincial senior high school entrance examination. Their mean score was 85.34 out of 120 (*SD = 9.86*), with the highest score being 99 and the lowest 71.

Prior to this study, all the participants and their English teacher were informed of the research purpose and that the results would be irrelevant to their evaluation at the school and their identities would remain confidential throughout the whole research process.

Instruments

On the basis of CSE (Ministry of Education of the People's Republic of China, 2018), which states that nine-graders roughly correspond to Level 3 and they should be able to "explain in simple terms the causes, processes, and results of events with generally accurate wording" and "use simple phrases to comment on familiar things and provide reasons, with generally coherent expression" (p.12), this study adopted the following two continuation tasks: Task 1 required students to read an incomplete story, in which the narrator was laughed at due to his limp, and then complete the story; while Task 2 required students to read a passage that introduced different sports popular in primary and middle schools, and then comment on sports. Both the input texts had been judged by the English teacher to be appropriate in terms of language and topic. Although the Chinese translations of the two texts were also provided to the students in case they didn't understand the input, Coh-Metrix L2 readability (McNamara et al., 2014) was calculated to be 27.375 and 22.95 for narrative and argumentative tasks respectively. One-sample T-test showed no significant difference ($t = 11.373, df = 1, p < 0.05$).

Data collection and analysis

The participants finished Task 1 first and then Task 2 a week later, as required by their English teacher, both within 30 min and with no less than 100 words. Then all the handwritten continuations were typed into electronic documents, and two small corpus were established, one containing 30 narrative continuations and the other one 30 argumentative ones.

In order to answer the above-mentioned three research questions, data analyses involved three steps. Firstly, all the 60 continuations were double-rated by two experienced English teachers according to the rubrics of NMET (Zhejiang version). The total inter-rater reliability was estimated to be 0.855 (0.905 and 0.847 for narrative and argumentative continuations respectively).

Secondly, the 60 continuations were analyzed in terms of various indices under the framework of CAF. Table 1 lists all the indices and the methods of how they were obtained. Table 2 lists the coding reliability of those manually-coded indices. Drawing on previous studies (e.g. Bi, 2020; Jiang & Chen, 2015; Lu, 2011; Shi et al., 2020; etc.), this study adopted nine and six indices to analyze syntactic and lexical complexity respectively; however, comparatively fewer indices were available to analyze accuracy and fluency, with only two and one respectively.

Thirdly, in order to find out the influence of genre on alignment, this study analyzed the frequency of 4-word phrases that were drawn from the input texts by students and LSS to measure alignment at syntactic and semantic levels. 4-word phrases were considered to be able to show how particular structures were borrowed

Table 1 Indices of CAF

Dimension	Indices	Methods
Complexity		
Syntactic	Mean Length of Sentence (MLS)	Coh-Metrix
	Mean Length of Clause (MLC)	Manual coding
	Mean Length of T-unit (MLT)	Manual coding
	Clause/T-unit (C/T)	Manual coding
	Dependent Clause/Clause (DC/T)	Manual coding
	Words before main verbs (WBV)	Coh-Metrix
	Number of modifiers before NP (NM)	Coh-Metrix
	Passive voice density (PD)	Coh-Metrix
	Syntax Similarity of adjacent sentences (SS)	Coh-Metrix
Lexical	Type-token ratio (TTR)	Coh-Metrix
	Measure of Textual Lexical Diversity (MTLD)	Coh-Metrix
	Incidence of content words (InCW)	Coh-Metrix
	Concreteness for content words (CCW)	Coh-Metrix
	Imageability for content words (ICW)	Coh-Metrix
	CELEX Log minimum frequency for content words (LogMF)	Coh-Metrix
Accuracy	error-free T-units/T-units (ET/T)	Manual coding
	errors/T-units (E/T)	Manual coding
Fluency	Words (W)	Coh-Metrix

Table 2 Coding reliability of manually-coded indices

Manually-coded indices	Narrative	Argumentative	Total
T-unit	0.974**	0.906**	0.968**
Error-free T-unit	0.903**	0.828**	0.908**
Clause	0.984**	0.953**	0.980*
Dependent Clause	0.871**	0.737**	0.809**
Error	0.737**	0.869**	0.795**

*Notes: *$p < 0.05$, **$p < 0.01$*

(Wang & Wang, 2014) and analyzed by means of the n-gram function of AntConc. Since alignment in written expressions is more associated with similar semantic network, rather than the borrowing of exact words (Wang & Wang, 2015), LSS was considered to be a more accurate indicator, which ranges from 0 to 1 (Zhang et al., 2023). The higher the LSS value is, the more similar the input text and the continuation are (Landauer & Dumais, 1997).

Results

Research question 1: influence of genre on scores

Table 3 lists the results of paired-samples T-test of students' scores of narrative and argumentative continuations. It can be seen that students performed only slightly better on narrative continuation (mean = 16.77) than on argumentative continuation

Table 3 Comparison of scores between genres

Genre	N	Min Score	Max Score	Mean Score	Std. deviation	Paired-samples T-test		
						t	df	*p*
Narrative	30	4.00	24.5	16.77	6.10	0.897	29	0.377
Argumentative	30	5.5	23.5	15.89	5.74			

Table 4 Comparison of syntactic complexity indices between genres

	Index	Mean	Std. deviation	t	df	P
1	MLS_NA-MLS_AR	-0.941	8.688	-0.593	29	0.558
2	MLC_NA-MLC_AR	-1.936	1.322	-8.022	29	0.000**
3	MLT_NA-MLT_AR	-2.511	1.443	-9.530	29	0.000**
4	WBV_NA-WBV_AR	-0.516	1.638	-1.724	29	0.095
5	NM_NA-NM_AR	-0.232	0.207	-6.123	29	0.000**
6	PD_NA-PD_AR	-0.477	7.712	-0.339	29	0.737
7	SS_NA-SS_AR	0.014	0.091	0.841	29	0.407
8	C/T_NA- C/T-AR	-0.048	0.104	-2.546	29	0.016
9	DC/T_NA- DC/T_AR	-0.076	0.118	-3.541	29	0.001**

*Note: **p < 0.01*

(mean = 15.89); however, this difference is not statistically significant ($t = 0.897$, $df = 29$, $p > 0.05$).

Research question 2: influence of genre on CAF indices

Influence of genre on syntactic complexity

Altogether there were nine indices used to show the influence of genre on syntactic complexity of students' continuations, five from Coh-metrix analysis (i.e. MLS, WBV, NM, PD and SS) and four through manual coding (i.e. MLC, MLT, C/T and DC/T). Table 4 lists the results of paired-samples T-test of these indices. Bonferroni adjustment were conducted to all *p*s to avoid false positive, making the critical value of significance 0.0056 (0.05/9) (Lu, 2010).

It can be seen that five out of the nine indices (i.e. MLS, WBV, PD, SS and C/T) did not show any difference between genres, while the other four indices (i.e. MLC, MLT, NM and DC/T) were significantly influenced by genre. To be specific, compared with argumentative continuations, narrative continuations contained much shorter clauses ($t = -8.022$, $df = 29$, $p < 0.0056$) and T units ($t = -9.53$, $df = 29$, $p < 0.0056$). Besides, in narrative continuations there were much fewer modifiers before noun phrases ($t = -6.123$, $df = 29$, $p < 0.0056$) and dependent clauses per T unit ($t = -3.541$, $df = 29$, $p < 0.0056$).

Influence of genre on lexical complexity

To compare lexical complexity, this study chose six indices from Coh-metrix analysis – TTR, MTLD, InCW, CCW, ICW and LogMF. Table 5 lists the results of paired-samples T-test of these indices. Again, Bonferroni adjustment was conducted by dividing 0.05 by 6 (i.e. 0.0083).

Table 5 Comparison of lexical complexity indices between genres

	Index	Mean	Std. deviation	t	df	P
1	TTR_NA-TTR_AR	-0.018	0.108	-0.927	29	0.362
2	MTLD_NA-MTLD_AR	0.177	32.795	0.030	29	0.977
3	InCW_NA-InCW_AR	-59.717	79.287	-4.125	29	0.000**
4	CCW_NA-CCW_AR	17.773	32.578	2.988	29	0.006**
5	ICW_NA-ICW_AR	13.479	29.613	2.493	29	0.019
6	LogMF_NA-LogMF_AR	0.304	0.392	4.249	29	0.000**

*Note: **p < 0.01*

It can be seen that half of the six indices showed significant influence of genre on lexical complexity, i.e. InCW, CCW, and LogMF. This means narrative continuations contained much fewer incidences of content words ($t = -4.125$, $df = 29$, $p < 0.0083$); but the concreteness ($t = 2.988$, $df = 29$, $p < 0.0083$) and CELEX Log minimum frequency ($t = 4.249$, $df = 29$, $p < 0.0083$) of these content words were much higher than those in argumentative continuations. These three indices demonstrate that students' narrative continuations were much easier to understand than argumentative ones.

Influence of genre on accuracy and fluency

This study used indices ET/T and E/T to measure the influence of genre on accuracy, and W on fluency. Table 6 lists the results of paired-samples T-test of these indices.

It can be seen that neither of the two indices of accuracy (ET/T and E/T) showed any significant difference ($ps > 0.05$), indicating that genre did not exert any influence on the accuracy of students' continuations. However, students' narrative continuations seemed to be much longer, with an average of 24.233 words more than argumentative ones, though this difference was not statistically significant enough ($p = 0.052$).

Table 6 Comparison of accuracy and fluency indices between genres

Index	Mean	Std. deviation	t	df	P
ET/T_NA-ET/T_AR	0.021	0.194	0.586	29	0.563
E/T_NA-E/T_AR	-0.055	0.265	-1.138	29	0.264
W_NA-W_AR	24.233	65.472	2.027	29	0.052

Table 7 Syntactic alignment in students' continuations

Frequency of 4-word phrases	Narrative	Argumentative	χ^2-test		
			χ^2	df	p
Aligned	7 (1.92%)	15 (7%)	9.568	1	0.002**
Non-aligned	358 (98.08%)	199 (93%)			
Total	365	214			

*Note: **p < 0.01*

Research question 3: influence of genre on alignment

Influence of genre on syntactic alignment

This study used the frequency of 4-word phrases that were drawn from the input texts to gauge the influence of genre on syntactic alignment in students' continuations. Table 7 lists the results of alignment of 4-word phrases.

It can be seen that students tended to borrow 4-word phrases more frequently when they were completing argumentative continuation task. To be specific, among all the 4-word phrases in narrative and argumentative continuations, students borrowed 7 and 15 times from the input texts, which accounts for 1.92% and 7% respectively of the total 4-word phrases. χ^2-test showed significant difference ($\chi^2 = 9.568$, $df = 1$, $p < 0.01$), indicating that there might be more syntactic alignment in the process of completing argumentative continuation task.

Influence of genre on semantic alignment

LSS analysis was conducted to investigate the degree to which students' continuations were semantically related to the input texts. Table 8 lists the results of LSS analysis, as well as the paired-samples T-test.

As shown in Table 8, compared with argumentative continuations, students' narrative continuations were more semantically similar to the input text, with an average LSS value of 0.59 ($t = 4.502$, $df = 29$, $p < 0.01$), suggesting that narrative continuation task might induce more semantic alignment.

Discussions

Examining the continuation task from the perspectives of both language testing and language learning, this study found that firstly, when scored holistically, students' continuations did not show significant difference between different genres of the input text; however, genre did have an impact on how students used the language to complete the continuation task, which could be shown by various indicators under the CAF framework on the one hand, and on the other hand by students' different ways of interacting with the input texts. These findings will be discussed in this section regarding the effects of genre of the input text on students' writing scores and language use.

Continuation task as a testing task: genre and writing scores

When used to assess students' EFL writing ability, continuation tasks in different genres were found to be able to bring about consistent performance measured by holistic score, similar to the results of Zhang et al. (2023), Zhong (2021) and Guo and Wang (2019).

Table 8 Semantic alignment in students' continuations

Genre	Sample	Min	Max	Mean	Std. deviation	Paired-samples T-test		
						t	df	p
Narrative	30	0.38	0.72	0.59	0.07	4.502	29	0.000**
Argumentative	30	0.23	0.77	0.47	0.13			

*Note: **p < 0.01*

These studies, involving students from middle school, high school and university and using continuation task as the testing task, seem to offer support for the shared-writing-ability conjecture (Kim & Crossley, 2018), which suggests that students' writing ability will not vary substantially under different testing conditions. Another possible explanation is that, despite the differences in overall English proficiency level and working memory capacity among the three age groups, the "scaffolding" function of the input texts may offset those differences to the extent that nuances in learners' continuations of different genres cannot be reflected by a holistic score. In addition, the data screening process in this study might be another cause for the non-significant result – those argumentative continuations that were off the point due to students' little experience with the continuation task were excluded from subsequent analysis, out of the consideration that the variable genre might get contaminated.

However, when continuation task is employed to test students' writing ability, it may be still premature to draw such a conclusion that genre does not matter, due to the limited number of studies in this regard. Like previous studies involving other types of writing tasks, genre may be interweaved with other factors like students' writing proficiency, education level, testing condition, topic, to name a few. The intricacies of genre-related test-methods effects require more investigation, especially with continuation tasks increasingly used in large-scale language tests.

Continuation task as a learning task: genre and language use

Despite that genre effects of the continuation task on holistically-scored writing quality may not be so obvious, as the present study shows, genre did influence students' use of language in a number of ways.

Firstly, guided by the CAF framework that is usually used to assess students' writing development, this study found that students' argumentative narrations contained more syntactically complex structures, as shown by the four indicators – MLC, MLT, NM and DC/T; besides, there were more content words in argumentative continuations, and they were much more abstract and appeared with a lower frequency in CELEX, making argumentative continuations more informative and more difficult to understand than narrative ones. Similar results about the comparative complexity of narrative and argumentative writings have been obtained in previous studies like Bi (2020), Lu (2011), Yoon and Polio (2016) and so on. However, the results of syntactic complexity in present study differed with those of Bi (2020), which also included students in middle school but only found significant difference in C/T and not in MLC and MLT. This might be due to both statistical and experimental reasons. Statistically, this study used Bonferroni adjustment when doing multiple comparisons, while Bi (2020) used $p = 0.05$ as the critical value. Experimentally, this study adopted repeated-measure design to control individual differences, while Bi (2020) compared the writings of two groups of students with similar writing proficiency.

As for the accuracy of students' continuations, results of different studies vary a lot. Some found no significant difference between two genres (e.g. this study; Guo & Wang, 2019; Yoon & Polio, 2016), some found students wrote more accurately in argumentative continuation (e.g.Leng, 2018; Zhang & Zhang, 2017) and some found the opposite (e.g. Yang, 2018b). These discrepancies indicate that how genre influences the accuracy

of students' continuations may be mediated by other factors, such as their education levels (junior high, senior high or college students), English proficiency levels (Elementary, Intermediate or Advanced stages) and indicators for accuracy (T unit-based indicators, error numbers of different types, or sub-score of language), and hence deserves more research efforts.

Similarly, the results about the genre effects on fluency are inconclusive, too. Compared with Zhang and Zhang (2017) that found students wrote much longer in narrative continuation, this study showed a tendency for narrative continuation (mean = 142.6) to be longer than argumentative one (mean = 118.4), but the difference was not statistically different ($p = 0.052$), probably because participants in the present study (nine graders) were not as proficient as those in Zhang and Zhang (2017) and hence not able to tell their stories freely to reveal the genre effect.

In a word, genre effects of the continuation task differ with dimensions of CAF, with complexity difference being the most obvious. As mentioned earlier, genre influences learners' continuations mainly through different demands of the limited working memory capacity, which has been found to be more significantly associated with CAF measures than overall writing proficiency (Li, 2023). In addition, the differences and similarities between genres in terms of CAF might partially support the Common Core Hypothesis (Leech & Svartvis, 2002), which suggests that when students complete multiple writing tasks in tandem, it is likely that their writings of different genres will share some characteristics. Therefore, as Zhang et al. (2022) commented, in order to account for L2 writing processes and products, it is desirable to "model both shared and distinct linguistic features of texts in different genres" (p. 7).

Secondly, this study found that different genres of continuation tasks may help students to interact with the input texts in different ways. To be specific, students borrowed more 4-word phrases from the input text in their argumentative continuations, suggesting that argumentative continuation task might bring about syntactic alignment. This is different from the results of Zhang and Zhang (2017) and Leng (2018). One possible reason is that they counted the number of aligned phrases manually and did not specify the length of phrases. Another reason might be that, since the English proficiency levels of participants in this study were comparatively lower, hence when they were completing argumentative continuation task, which might be more cognitively demanding for their age, they had to resort to the input text and borrow phrases in the writing process.

In addition, students' narrative continuations were found to be more semantically associated with the input text, suggesting a stronger lexical alignment of narrative continuation task. This is consistent with the results of Zhang et al. (2023), probably because middle school students have already been familiar with the narrative genre and are able to spare some cognitive capacity to continue the story with language consistent with their own proficiencies. However, the LSS values of this study were much lower, probably because leaners in this study were less trained in how to complete continuation task.

Conclusion

This study examined the influence of input text genre on Chinese young EFL learners' performance on the continuation task. Taking continuation task as a type of test task, this study found that genre did not influence students' writing scores, offering another

piece of validity evidence for using it in various testing programs. Taking it as a learning task, this study revealed varying genre effects on the dimensions of complexity, accuracy and fluency, and also on how students interacted with the input text syntactically and semantically, suggesting that teachers may use input texts in different genres to achieve their goals of training.

There are some limitations in this study. First, with only two continuation tasks, one narrative and one argumentative, used to compare students' performance, there might be some measurement errors. Future research could increase the number of tasks to assess students' performance more accurately. Second, in the comparison of linguistic features, like many previous studies, this study used much more indicators of complexity than indicators of accuracy and fluency. This will cause trouble not only for cross-study comparisons of complexity, but also for the accurate measurement of accuracy and fluency. Future studies could find out the most influential complexity indicators on one hand and develop more tools to facilitate the analyses of accuracy and fluency. Thirdly, although the two input texts were evaluated by the learners' English teacher in terms of language and topic, more direct evidence should have been collected from learners themselves, so as to eliminate the possible influence of factors other than genre.

Abbreviations

CAF	Complexity, accuracy and fluency
CCW	Concreteness for Content Words
CSE	China's Standards of English Language Ability
C/T	Clause per T-unit
DC/T	Dependent Clause per T-unit
EFL	English as a Foreign Language
E/T	Errors per T-unit
ET/T	Error-free T-units per T-unit
ICW	Imageability for Content Words
InCW	Incidence of Content Words
LogMF	CELEX Log Minimum Frequency for Content Words
LSM	Language Style Matching
LSS	Latent Semantic Similarity
MLC	Mean Length of Clause
MLS	Mean Length of Sentence
MLT	Mean Length of T-unit
MTLD	Measure of Textual Lexical Diversity
NMET	National Matriculation Entrance Test
NM	Number of Modifiers before Noun Phrase
PD	Passive Voice Density
SS	Syntactic Similarity of Adjacent Sentences
TTR	Type Token Ratio
WBV	Words before Main Verbs

Availability of data and materials
The datasets used and / or analyzed during the current study are available from the author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Berman, R. A. (2008). The psycholinguistics of developing text construction. *Journal of Child Language*, *35*(4), 735–771. https://doi.org/10.1017/S0305000908008787

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, *43*(2), 79–120. https://doi.org/10.1207/s15326950dp4302_1

Bi, P. (2020). Revisiting genre effects on linguistic features of L2 writing: A usage-based perspective. *International Journal of Applied Linguistics*, *30*(3), 429–444. https://doi.org/10.1111/ijal.12297

Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly*, *48*(4), 738–762. https://doi.org/10.1002/tesq.148

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Grabe, W. (2002). Narrative and expository macro-genres. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 249–267). L. Erlbaum.

Gu, Q., Zhao, C., & Jin, X. (2022). Effects of topic familarity on continuation task among Chinese EFL learners. *Journal of PLA University of Foreign Languages*, *45*(2), 43–51.

Guo, Y., & Wang, X. (2019). Effects of the reading material on English continuation. *Journal of Donghua University (Social Science)*, *19*(4), 425–435. https://doi.org/10.19883/j.1009-9034.2019.04.014

Jiang, L., & Chen, J. (2015). The continuation task: Effects on written accuracy, complexity and fluency. *Modern Foreign Languages*, *38*(3), 366–375.

Jiang, L., & Tu, M. (2016). Effects of the continuation task on L2 vocabulary learning. *Modern Foreign Languages*, *39*(6), 819–829.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Lawrence Erlbaum Associates.

Kim, M., & Crossley, S. A. (2018). Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing*, *37*, 39–56. https://doi.org/10.1016/j.asw.2018.03.002

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisiton, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. https://doi.org/10.1037/0033-295X.104.2.211

Leech, G., & Svartvis, J. (2002). *A communicative grammar of English*. Routledge.

Leng, C. (2018). *Genre-related alignment in English continuation task on junior high*. MA Thesis, Chongqing Normal University.

Li, S. (2023). Working memory and second language writing: A systematic review. *Studies in Second Language Acquisition*, *45*(3), 647–679. https://doi.org/10.1017/s0272263123000189

Li, J. (2014). Rasch analysis of summary writing tasks of different genres. *Foreign Languages and Their Teaching*, (5), 30–35. https://doi.org/10.13458/j.cnki.flatt.004017

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

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

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

Ministry of Education of the People's Republic of China. (2018). *China's standards of English language ability (GF 0018)*. Higher Education Press.

Peng, H. (2017). An empirical study on English learners' coherence in writing. *Journal of PLA University of Foreign Languages*, *40*(4), 87–92.

Peng, J., Wang, C., & Lu, X. (2020). Effect of the linguistic complexity of the input text on alignment, writing fluency, and writing accuracy in the continuation task. *Language Teaching Research*, *24*(3), 364–381. https://doi.org/10.1177/1362168818783341

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–226. https://doi.org/10.1017/S0140525X04000056

Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing*, *33*, 3–17. https://doi.org/10.1016/j.jslw.2016.06.001

Ren, W., & Lv, X. (2021). A meta-analysis of the alignment of continuation writing tasks. *Foreign Language World*, (4), 44–52.

Shao, J. (2021). MFRM-based scoring validity of self-chosen-task writng test with multiple genres. *Journal of Xi'an International Studies University*, *29*(3), 72–77. https://doi.org/10.16362/j.cnki.cn61-1457/h.2021.03.015

Shi, B., Huang, L., & Lu, X. (2020). Effect of prompt type on test-takers' writing performance and writing strategy use in the continuation task. *Language Testing*, *37*(3), 361–388. https://doi.org/10.1177/0265532220911626

Sun, Q., & Wang, Y. (2018). Effectiveness of continuation tasks on the learning of English articles. *Technology Enhanced Foreign Langauge Education*, (2), 18–24.

Sun, X., & Wang, Y. (2022). A study of the effect of the continuation task on Chinese EFL learners' article use. *Foreign language and Their Teaching*, (5), 87–94＋147–148. https://doi.org/10.13458/j.cnki.flatt.004893

Swales, J. M. (1981). *Aspects of article introductions: language studies unit*. University of Aston.

Tardy, C. M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of Second Language Writing*, *15*(2), 79–101. https://doi.org/10.1016/j.jslw.2006.04.003

Wang, Q., & Wang, C. (2019). Learning English relative clauses by extension. *Foreign Language Learning Theory and Practice*, (3), 1–5＋18.

Wang, Q., & Cao, Q. (2020). Structural priming in L2 continuation: Take English passive sentences as an example. *Journal of PLA University of Foreign Languages*, *43*, 25–32.

Wang, M., & Wang, C. (2014). The effects of alignment in continuation tasks. *Modern Foreign Languages*, *37*, 501–512＋584.

Wang, C., & Wang, M. (2015). Effect of alignment on L2 written production. *Applied Linguistics*, *36*(5), 503–526. https://doi.org/10.1093/applin/amt051

Wang, H., Yang, Y., & Yang, Z. (2022). Interaction between linguistic complexity and accuracy of L2 writing: A cross-genre analysis. *Foreign Language Learning Theory and Practice*, (02), 39–48＋162.

Wang, C. (2012). Continuation: An efficient way to improve foreign language learning. *Foreign Language World*, (5), 2–7.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawai'i Press.

Xiao, T. (2013). *Effect of alignment on improving L2 accuracy*. MA Thesis, Guangdong University of Foreign Studies.

Xin, S. (2017). Effects of continuation task condition on L2 grammatical structure acquisition. *Modern Foreign Languages*, *40*, 507–517.

Xin, S., & Li, L. (2020). Alignment of textual complexity in continuation task and its effects on accuracy. *Journal of PLA University of Foreign Languages*, *43*, 33–41.

Xue, H. (2013). *Interest-related alignment in L2 story continuation*. [MA dissertation] Guangdong University of Foreign Studies.

Yang, H. (2018a). Learning effects of extension writing on the written rhetoric of higher intermediate foreign language learners. *Foreign Language Teaching and Research*, *50*(4), 596–607.

Yang, X. (2018b). *The effect of text genre on alignment in continuation task*. MA Thesis, Sichuan Normal University.

Ye, Y., & Yan, P. (2010). Influence of genre variation on writing performance. *Foreign Language World*, (2), 63–70.

Yoon, H.-J., & Polio, C. (2016). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, *51*(2), 275–301. https://doi.org/10.1002/tesq.296

Yoon, H.-J. (2021). Challenging the connection between task perceptions and language use in L2 writing: Genre, cognitive task complexity, and linguistic complexity. *Journal of Second Language Writing, 54*. https://doi.org/10.1016/j.jslw.2021.100857

Zhang, L., & Liu, H. (2021). Genre effect on L2 syntactic complexity and holistic rating for writing quality of intermediate EFL learners. *Chinese Journal of Applied Linguistics*, *44*(04), 451–469＋589.

Zhang, L., & Qin, T. (2020). Effects of the continuation task on EFL learners' writing anxiety and writing ability. *Foreign Language Teaching*, *41*(6), 72–76. https://doi.org/10.16362/j.cnki.cn61-1023/h.2020.06.014

Zhang, X., Lu, X., & Li, W. (2022). Beyond differences: Assessing effects of shared linguistic features on L2 writing quality of two genres. *Applied Linguistics*, *43*(1), 168–195. https://doi.org/10.1093/applin/amab007

Zhang, J., Wang, M., & Chen, K. (2023). Effects of source text genre on linguistic alignment and writing quality in continuation tasks. *Modern Foreign Languages*, *46*(2), 259–269.

Zhang, X., & Wang, M. (2022). The continuation task and the model-as-feedback writing task in L2 vocabulary learning: Timing of model texts. *Language Teachching Research*. https://doi.org/10.1177/13621688221119077

Zhang, X., & Zhang, Q. (2017). Influence of continuation task in different genres on alignment effect. *Foreign Language World*, (3), 90–96.

Zhang, X. (2016). Effects of the continuation task on ESL writing process. *Foreign Language World*, (6), 86–94.

Zhong, J. (2021). *Investigating the influence of genre type on students' English writing performance in the continuation task*. PhD Thesis, Queen's University.

Zhu, Z., & Cao, Y. (2020). Scaling descriptors for China's standards of English langauge ability. *Foreign Languages in China*, *17*(4), 15–22. https://doi.org/10.13564/j.cnki.issn.1672-9382.2020.04.003

Publisher's Note