
| RESEARCH ARTICLE

Effects of Task Type on Linguistic Complexity of Chinese EFL Undergraduates: A Functional Linguistics Perspective

Keqing Yu

School of Foreign Languages, Guangzhou College of Commerce, Guangzhou, China

Corresponding Author: Keqing Yu, **E-mail:** yukeqing036@outlook.com

| ABSTRACT

Although numerous studies have explored the relationship between linguistic complexity and L2 writing quality, most have focused on argumentative or persuasive writing, leaving uncertainties about the potential genre effects on linguistic complexity variation. Therefore, this study aimed to investigate how different task types, based on a functional linguistics taxonomy, influence linguistic complexity. Intermediate-to-high proficiency EFL university students (N = 144) completed argumentative, descriptive, interactive, and narrative essays. The results revealed that students performed best on argumentative tasks, with significant differences observed in lexical complexity between argumentative and descriptive essays, and in syntactic complexity between argumentative/descriptive and interactive/narrative essays. Lexical complexity emerged as the strongest predictor of writing scores across all four task types. Pedagogical implications are discussed.

| KEYWORDS

Task type, lexical complexity, syntactic complexity, EFL learners

| ARTICLE INFORMATION

ACCEPTED: 01 November 2024

PUBLISHED: 21 November 2025

DOI: 10.32996/ijllt.2025.8.12.1

1. Introduction

Language use, as an essential component of second language (L2) assessment criteria, has attracted significant attention from numerous L2 writing researchers. Various linguistic features have been explored, with particular emphasis on linguistic complexity as a predictor of L2 writing quality within the domain of language testing and assessments (e.g., Barrot & Agdeppa, 2021; Kim & Crossley, 2018; Peng et al., 2023). Likewise, linguistic complexity has played a pivotal role in tracking L2 language development in studies on language acquisition (e.g., Kuiken, 2023; Polat & Kim, 2014) and in comparing complexity variation across different classroom-based teaching contexts (e.g., Fathi & Rahimi, 2022).

Linguistic complexity, conceptualized as a multi-dimensional construct, typically encompasses both lexical and syntactic complexity. Within the context of L2 writing assessments, prior research has demonstrated that more proficient writers are capable of producing more complex syntactic structures and utilizing vocabulary more effectively. In L2 acquisition studies, lexical and syntactic complexity serve as key indicators of linguistic performance, exhibiting non-linear development across different stages of language acquisition (e.g., Lambert & Kormos, 2014). In the field of L2 language teaching, both lexical use and global-level syntactic complexity have been found to be particularly sensitive to linguistic development (e.g., Fathi & Rahimi, 2022). However, it is important to recognize the interplay between writing genres and linguistic complexity. Specifically, the manifestation of lexical and syntactic complexity can vary considerably across different writing task types (e.g., Latif & Gierl, 2019; Pu, Heng, & Cao, 2022; Zhen & Barrot, 2024; Zheng & Barrot, 2024). Research comparing independent and integrated writing tasks has shown that L2 writers tend to produce more sophisticated vocabulary and complex syntactic structures when external input is provided, as this external scaffolding stimulates their linguistic repertoire (Kyle & Crossley, 2016; Tabari, Lu, & Wang, 2023). Nonetheless, these findings are largely based on argumentative writing, which shares common generic processes such as taking and defending a position.

The Systemic-Functional Linguistics (SFL) taxonomy of task types has been widely applied in large-scale, high-stakes English proficiency tests, such as the IELTS. Adopting a functional linguistics perspective, writing tasks can be categorized into argumentative (i.e., arguing for a stance), descriptive (i.e., describing given information), expository (i.e., explaining or illuminating written materials), narrative (i.e., narrating an event), interactive (i.e., engaging in interpersonal relationships), and instructive (i.e., instructing or informing) tasks (Halliday & Hasan, 1989; Halliday & Matthiessen, 2013). While there is increasing interest in the variation of linguistic complexity across different task types, most existing studies have focused on binary comparisons (e.g., Barrot & Agdeppa, 2021; Peng et al., 2023; Zhang & Li, 2023) or on a single sub-construct of linguistic complexity (e.g., Khushik & Huhta, 2020; Zheng & Barrot, 2024). As a result, there remains a significant gap in research regarding how linguistic complexity varies across various writing tasks and which linguistic features predict L2 writing quality across genres. The present study aims to address these gaps by examining how lexical and syntactic complexity indices differ across argumentative, descriptive, interactive, and narrative writing tasks. The findings from this research will enhance our understanding of genre effects on linguistic performance and contribute to a more nuanced view of linguistic complexity. From a pedagogical standpoint, these findings will provide language teachers with valuable insights into students' writing performance in diverse task contexts.

2. Literature Review

2.1 Studies on lexical complexity in L2 writing

Lexical complexity, a multi-faceted construct, is defined as “a wide variety of basic and sophisticated words that are available and can be accessed quickly” (Wolfe-Quintero, Inagaki, & Kim, 1998, p. 102). It includes three sub-constructs: lexical sophistication, lexical diversity, and lexical density (Housen & Kuiken, 2012; Tabari, Lu, & Wang, 2023). This study focuses on two major components: lexical sophistication and lexical diversity.

Kyle and Crossley (2015) distinguished lexical sophistication into two major aspects: word frequency and psycholinguistic information, which can be measured using natural language processing (NLP) tools such as Coh-Metrix (Graesser, McNamara, & Kulikowich, 2011). Word frequency reflects “the probability at which a word is encountered by L2 learners in language input” (Peng et al., 2023, p. 3). Psycholinguistic information assesses L2 learners' word knowledge (Salsbury et al., 2011). Lexical sophistication has been shown to be an indicator of L2 writing quality. For instance, Kim and Crossley (2018) investigated the contribution of lexical features to TOEFL iBT independent and integrated essay scores. They found that one measure of lexical complexity, lexical decision mean response times ($\beta = 0.932$, $p < 0.01$), exhibited strong correlations with quality ratings. Similarly, Tabari et al. (2023) measured lexical sophistication using three TAALES-based indices (i.e., COCA Log Frequency AW, COCA Log Frequency CW, and COCA Log Frequency FW) and found them sensitive to variations in L2 writing quality, performing consistently across both simple and complex argumentative essays.

Lexical diversity, by contrast, pertains to “the range of unique words (types) that occur in a given text” (Graesser, 2011, p. 226) and measures the breadth of vocabulary knowledge (Crossley et al., 2011; Peng et al., 2023). Research on lexical diversity has received less attention, with its relationship to L2 writing quality being relatively weak. For example, Latifi and Gierl (2021), using a sample size of $N = 12,978$, examined the contribution of linguistic features to overall writing scores. Their findings indicated that lexical diversity accounted for 5.3% of the total score variance, compared to 15.2% from word-frequency-based lexical complexity indices. However, it is important to note that the contributions of lexical complexity features vary across different task types, ranging from 3.7% to 6.5%, indicating a strong influence of genre on lexical complexity variation.

2.2 Studies on syntactic complexity and L2 writing quality

Syntactic complexity, another key dimension of linguistic complexity, refers to an L2 writer's ability to produce diverse and intricate syntactic structures (Ortega, 2003; Wolfe-Quintero et al., 1998). This dimension has been the focus of numerous studies aimed at differentiating learners at various proficiency levels, predicting L2 writing quality, and examining language development. In the design of L2SCA (Second Language Syntactic Complexity Analyzer), Lu (2010) adopted Ortega's (2003) definition and operationalized syntactic complexity using fourteen measures across five levels: length of production units, amount of subordination, amount of coordination, degree of phrasal sophistication, and overall sentence complexity (Lu, 2011; Lu & Ai, 2015). These measures include: mean length of sentence (MLS), mean length of T-unit (MLT), mean length of clause (MLC), clauses per T-unit (C/T), complex T-units ratio (CT/T), dependent clauses per clause (DC/C), dependent clauses per T-unit (DC/T), coordinate phrases per clause (CP/C), coordinate phrases per T-unit (CP/T), T-units per sentence (T/S), complex nominals per clause (CN/C), complex nominals per T-unit (CN/T), verb phrase per T-unit (VP/T), and clauses per sentence (C/S). These indices have been proven effective in predicting changes in L2 writing quality (e.g., Abrams, 2019; Barrot & Agdeppa, 2021; Khushik & Huhta, 2020; Plakans, Gebril, & Bilki, 2019; Yang, Lu, & Weigle, 2015; Zheng & Barrot, 2024). For example, Plakans et al. (2019) found that mean T-unit length in TOEFL iBT integrated essays ($N = 480$) increased linearly as writing scores improved. Yang et al. (2019) studied syntactic complexity variation across different topics among EFL adult learners ($N = 386$) and found that two global indices—mean sentence length and mean T-unit length—were strong predictors of writing scores, while local-level syntactic complexity measures, such as DC/T, displayed topic effects. Furthermore, a recent study by Zheng and Barrot (2024) investigated task type effects on syntactic complexity using narrative and argumentative essays from college students (N

= 161). Surprisingly, the study found that argumentative essays outperformed narrative essays in six syntactic complexity measures (MLT, MLC, CP/C, CP/T, CN/C, and CN/T), suggesting a potential interaction between genre and syntactic complexity.

2.3 Genre effects on linguistic complexity in L2 writing

Numerous studies on independent (i.e., tasks requiring test-takers to respond based on prior knowledge) and integrated (i.e., tasks requiring test-takers to integrate given materials) writing tasks have focused on how these tasks differ in terms of linguistic features. Research has shown that EFL writers tend to perform better in integrated writing tasks. For instance, Kim and Crossley (2018), using NLP tools, examined lexical, syntactic, and cohesive differences between TOEFL iBT independent and integrated essays. The study found that three TAALES-based lexical sophistication measures (lexical decision mean reaction time, word naming mean reaction time, and age-of-acquisition) and one SCA-based syntactic complexity measure (mean length of clause) were significantly higher in integrated essays, suggesting that integrated tasks elicit more complex syntactic structures and more advanced vocabulary. A similar pattern was observed in a study by Riazi (2016), who found that lexical sophistication was a strong predictor of both independent and integrated writing scores. However, it is important to note that these two tasks share certain generic processes, such as argumentation, which may account for their similarities in specific linguistic features.

Research into the impact of task type on linguistic performance has garnered significant interest from L2 writing researchers, particularly regarding argumentative, expository, descriptive, and narrative writing within the SFL taxonomy. Argumentative, expository, and descriptive tasks are commonly used to assess test-takers' academic writing abilities, while narrative writing is primarily employed with younger or low-proficiency learners (Fitzgerald & Spiegel, 1986; Tabari & Johnson, 2023).

Most existing studies on genre effects in L2 writing compare linguistic performances across two genres (e.g., Jeong, 2017; Pu, Heng, & Cao, 2022; Zhang & Li, 2023; Zhang, Lu, & Li, 2022; Zheng & Barrot, 2024). For instance, Zhang and Li (2023) compared lexical and syntactic complexity in EFL learners' (N = 614) argumentative and expository essays. They found that all seven lexical sophistication and seven syntactic complexity indices were higher in argumentative essays. Similar findings were reported by Pu et al. (2022), who specifically examined syntactic complexity variations across argumentative and expository tasks. It was found that argumentative essays outperformed expository essays in two global-level (MLS and C/S), three clausal-level (CT, DC/C, and DC/T), and one phrasal-level (VP/T) indices. A similar pattern was observed when comparing argumentative and narrative essays in the study by Zheng and Barrot (2024). However, narrative and expository essays showed significant differences in lexical use. Jeong (2017) examined the role of genre (narrative vs. expository) in performance criteria, using data from EFL adult learners (N = 180). The results revealed that expository writing scores were more centralized than those for narratives, and students used a higher range of vocabulary in expository writing. Moreover, systemic differences were also found between interactive and argumentative writing (e.g., Zhang, Lu, & Li, 2022).

While these studies provide valuable insights into genre effects on L2 writing quality, research covering more task types with different generic processes has been relatively scarce (e.g., Latif & Gierl, 2021). In one study, Latif and Gierl (2021) explored how linguistic features varied across persuasive, expository, and narrative tasks. They found that lexical complexity indices were more indicative of narrative scores, while syntactic complexity indices were least indicative of these scores. However, the lack of detailed within-group information warrants further exploration. Thus, research examining systemic task-type differences in lexical and syntactic complexity is necessary.

Taken together, the studies reviewed either focused exclusively on argumentative writing (e.g., Kim & Crossley, 2018; Yang et al., 2019), on a single dimension of linguistic complexity (e.g., Khushik & Huhta, 2020; Zheng & Barrot, 2024), on comparisons between two task types (e.g., Zhang et al., 2022; Zheng & Barrot, 2024), or lacked detailed information (e.g., Latif & Gierl, 2021).

Therefore, the present study aims to examine how task types with distinct generic processes (i.e., argumentation, narration, description, and interaction) affect linguistic complexity. This study is guided by the following three research questions:

RQ1: How do different task types (argumentative writing, descriptive writing, interactive writing, and narrative writing) affect the lexical and syntactic complexity of Chinese undergraduates?

RQ2: What is the predicting power of lexical and syntactic complexity measures to writing scores across task types?

3. Methodology

3.1 Participants

The participants in this study were 36 EFL undergraduates majoring in English at a Chinese university, all of whom voluntarily participated. They were enrolled in seven classes within the same grade. On average, participants had received thirteen years of English education and over three and a half years of professional English immersion courses. Based on alignment with the China Standards of English (CSE) (Peng, Liu, & Cai, 2021), the overall proficiency of all participants ranged from level 4 to level 7, which corresponds roughly to CEFR levels A2 to B2.

Two raters, both experienced English teachers with master's degrees in linguistics, were recruited for this study. They each had at least sixteen years of English learning experience and over two years of rater training.

3.2 Writing Tasks

This study included four writing tasks: argumentative, descriptive, interactive, and narrative, all of which are commonly used in Chinese classroom instruction and official proficiency tests. The tasks were randomly selected from the 2021-2023 College English Test-Band 6 (CET-6) writing sections (Zheng & Cheng, 2008), as previous research has validated the comparability of these tasks with different topics (e.g., Miao, 2006), thus eliminating potential topic effects.

The argumentative writing (AW) task required participants to take a stance on a topic and support it. In the descriptive writing (DW) task, a line chart was provided, and participants were required to describe and briefly comment on the information presented in the chart. The interactive writing (IW) task involved writing a proposal for school services, while the narrative writing (NW) task required participants to narrate their voluntary visit to an elementary school or nursing home. All materials were printed to ensure a quasi-experimental environment. The scoring rubrics for these tasks included topic relevance, cohesion and coherence, linguistic accuracy, and complexity.

3.3 NLP tools

To measure syntactic complexity, the Second Language Syntactic Complexity Analyzer (L2SCA), developed by Lu (2010, 2015), was used. This tool is widely used for the automatic analysis of syntactic complexity in written English. It includes fourteen indices that measure global, clausal, and phrasal dimensions of syntactic complexity. The developers of L2SCA reported high reliability for the more complex production units and syntactic structures, with correlations between complexity scores and the fourteen indices ranging from 0.834 to 1.000 (Ai & Lu, 2013; Lu & Ai, 2015). Due to the unavailability of the website-based L2SCA, NeoSCA, a fork of L2SCA with added support for Windows, was used in this study. Syntactic complexity was measured by the fourteen indices provided by NeoSCA (see Table 1).

Coh-Metrix 3.0, a computational tool for measuring lexical and syntactic complexity, cohesion, and coherence, was used to measure lexical complexity in this study. Lexical complexity was assessed using two word frequency indices (word frequency and minimum word frequency per sentence), five psycholinguistic complexity indices (word familiarity, imageability, meaningfulness, polysemy, and hypernymy), and four lexical diversity indices (type-token ratio, lexical diversity for all words, textual lexical diversity, and VOC lexical diversity) (see Table 1).

3.4 Procedures and data collection

All participants signed a consent form and gave approval for the use of their data. They completed two writing tasks at a time, with thirty minutes allotted for each task. The use of electronic devices or dictionaries was prohibited. The paper-and-pencil tasks were administered in a quasi-test environment, with assistance from a student assistant. After completing the four essays, participants filled out a linguistic background questionnaire.

All essays were directly scored by the two raters and then transcribed into electronic text using Microsoft Word for later analysis. Spelling errors were manually corrected because Coh-Metrix and NeoSCA are sensitive to misspellings. The corrected texts were then input into the NLP tools for analysis. Before scoring, the raters discussed the scoring descriptors. During the rating process, negotiations took place whenever their scores for a single essay differed by more than three points. To ensure inter-rater reliability, Cronbach's alpha was computed, yielding a coefficient of $\alpha = .838$.

3.5 Data analysis

No significant differences in writing scores were found between the four tasks. Normality tests were conducted on the fourteen syntactic complexity and eleven lexical complexity indices across the four writing tasks. To address RQ1, the normally distributed data were analyzed using homogeneity tests and repeated-measure ANOVA to examine within-group differences. For data that violated normality assumptions, Friedman tests were conducted. Post-hoc tests were performed to identify specific within-group differences.

To address RQ2, stepwise multiple linear regression analyses were conducted to determine the best predictors of writing scores across the four task types. Lexical and syntactic complexity indices were used as independent variables, and writing scores were treated as dependent variables in SPSS 23.0.

4. Results

4.1 Differences in lexical complexity indices across four task types

Table 2 shows means and standard deviations of lexical complexity variation on four writing tasks. First, argumentative, descriptive, interactive, and narrative essays performed alike in two lexical diversity indices (LDTRc and LDTRa), all word frequency indices (WRDFRQa and WRDFRQmc), and one psycholinguistic index (WRDHYPn). Notably, LDTRc and LDTRa exhibit consistently low mean values across all tasks, with IW demonstrating the highest mean for both metrics, suggesting a slight increase in lexical diversity for this task. Second, there is a drastic difference between descriptive writing and the other three. LDMTLD ($M = 61.56$, $SD = 14.68$), LDVOC ($M = 59.47$, $SD = 18.51$), WRDMEAc ($M = 442.57$, $SD = 14.68$), and WRDPOLc ($M = 3.94$, $SD = .52$) were the lowest in descriptive essays. In contrast, LDMTLD and LDVOC display significantly higher mean values, particularly in DW and AW tasks, respectively, showcasing greater lexical diversity and vocabulary richness. Third,

argumentative essays displayed the lowest values of WRDIMGc ($M = 386.41$, $SD = 21.06$). It should be also noted that all four lexical diversity indices, WRDFAMc, and WRDPOLc in descriptive writing were found relatively lower than those of other tasks. Table 3 displays within-group differences in lexical complexity between four task types. It is obvious that three lexical diversity indices (LDTTRc, LDMTLD, and LDVOCd) were significantly lower in descriptive essays ($M[LDTTRc] = .69$, $SD = .07$, $p[DW, NW] < .05$; $M[LDMTLD] = 61.56$, $SD = 14.68$, $p[AW, DW] < .05$; $M[LDVOCd] = 59.47$, $SD = 18.51$, $p[DW, AW/IW/NW] < .05$). In terms of word frequency, argumentative and descriptive essays showed significantly lower WRDFRQmc compared to interactives and narratives. In terms of psycholinguistic complexity, there was a significant improvement of WRDIMGc in narratives, followed by interactive, descriptive, and argumentative performances. Besides, there was a significant decrease in WRDFAMc in descriptive writing ($M[WRDFAMc] = 442.57$, $SD = 14.68$, $p[DW, AW/IW/NW] < .05$).

4.2 Differences in syntactic complexity indices across four task types

Table 4 exhibits means and standard deviations of syntactic complexity indices across four tasks. First, DW registered higher three global indices (MLS, MLT, and MLC; $M[MLS] = 17.5$, $SD = 4.58$; $M[MLT] = 16.68$, $SD = 4.44$; $M[MLC] = 16.68$, $SD = 2.41$) and two phrasal indices (CN/C and CN/T; $M[CN/C] = 1.48$, $SD = .44$; $M[CN/T] = 2.22$, $SD = .78$), while AW showed improvement in the majority of clausal complexity indices (i.e., C/S, C/T, DC/C, DC/T, T/S, and CT/T). Moreover, global syntactic complexity indices (MLS, MLT, and MLC) were found relatively lower in IW ($M[MLS] = 15.64$, $SD = 3.85$; $M[MLT] = 14.77$, $SD = 3.65$; $M[MLC] = 14.77$, $SD = 2.05$), suggesting a simpler structure in IW. It was also found that all clausal-level indices (C/S, C/T, DC/C, DC/T, T/S, and CT/T) in NW registered the lowest frequency.

Table 5 shows how syntactic complexity indices varied across four tasks. Overall, smaller task type differences were observed. First, one global (C/S) and three clausal (DC/C, DC/T, and CT/T) significantly varied between AW and NW. DW and NW varied significantly solely in CN/T ($p < .05$). AW and DW only differed significantly in MLC and CN/C. Two phrasal indices (CN/T and CN/C) showed significant variation between IW and DW (interactive: $M[CN/T] = 1.58$, $SD[CN/T] = .54$, $M[CN/C] = 1.08$, $SD[CN/C] = .32$, $p[IW, DW] < .05$; descriptive: $M[CN/T] = 2.22$, $SD[CN/T] = .78$, $M[CN/C] = 1.48$, $SD[CN/C] = .44$, $p[IW, DW] < .05$).

4.3 Lexical and syntactic complexity indicators of writing scores

Table 6 summarizes the final regression results (Model 4). The final regression model explained 26.2% of the total score variance. LDTTRc made the largest contribution, followed by WRDIMGc, WRDMEAc, and WRDFRQa. VIF values were acceptable, indicating no violation of multi-collinearity. Moreover, LDTTRc ($B = 12.355$, $SE = 2.287$) and WRDIMGc ($B = .031$, $SE = .010$) positively correlated overall essay ratings, while WRDMEAc and WRDFRQa showed negative relationships ($B[WRDMEAc] = -.033$, $SE = .015$; $B[WRDFRQa] = -3.944$, $SE = 1.805$). The results indicate that higher-rated essays tended to contain more sophisticated vocabulary.

5. Discussion

The first RQ is associated with how lexical complexity varied across argumentative, descriptive, interactive, and narrative writing tasks. In our analyses, five indices showed significant differences between descriptive and other essays. These include three indices of lexical diversity, namely, LDTTRc, LDMTLD, and LDVOCd, and two indices of psycholinguistic complexity, namely, WRDFRQmc and WRDFAMc. It suggests that the students used some words multiple times in the descriptive task. Additionally, argumentative essays registered lower two indices of psycholinguistic complexity, i.e., WRDIMGc and WRDMEAc. These results suggest that the learners tended to utilize more sophisticated and diverse vocabulary in argumentative writing, followed by interactives and narratives. This finding partly aligns with previous studies (Zhang et al., 2022) that found L2 writers outperformed on argumentative tasks. Second, narratives showed higher psycholinguistic indices, namely, WRDIMGc and WRDMEAc, indicating higher word repetitions and lower lexical sophistication. This finding partly echoes the results of Jeong (2017) who found narrative texts reported lower scores of vocabulary use.

The second research question concerned the role of four task types in syntactic complexity. In global and sentential complexity, argumentative texts outperformed descriptive texts in MLC and narrative texts in C/S. The observed task type differences in the length of production unit could be associated with the cognitive demands of different writing tasks. Descriptive writing in this study requiring the writers to describe a line chart and then argue for a viewpoint seemed to involve not just the processes of description and argumentation, but reading and writing abilities. As evidenced by Yang (2016), descriptive (graph) writing ability significantly correlated graph interpretation ability and linguistic competence. Second, in terms of clausal complexity, three subordination indices (i.e., DC/C, DC/T, and CT/T) in argumentative essays showed significant variation from those in narratives. Some studies have reported writing proficiency-based variation in subordination (e.g., Ortega, 2003; Zhang et al., 2022), but often with statistical insignificance in relationships with writing scores. One possible reason was that narrative writing has been mostly used and practiced for young learners (Tabari & Johnson, 2023) and classroom-based teaching focuses on argumentative writing because the latter has been validated for its close relationship with academic writing ability (Shin & Ewert, 2015). In terms of phrasal complexity, there was a significant improvement of CN/T and CN/C in descriptive writing, indicating that they may be a robust tool for differentiating descriptive and other three tasks.

The regression analyses revealed that lexical complexity indices were predictive of overall essay ratings. One index of lexical diversity, LDTRc, was the strongest predictor, suggesting that higher-rated essays tended to contain more presences of unique words. Such findings were consistent with previous studies (Latif & Gierl, 2019; Lee, Ge, & Chung, 2021) which explored the predicting power of linguistic complexity features for writing scores. Two psycholinguistic complexity indices (WRDIMGc and WRDMEAc) showed opposite relationships with writing scores. One index of word frequency (WRDFRQa) negatively correlated overall essay ratings. Similar results were found in Kyle and Crossley (2016).

Not surprisingly, similar to the results of Lee et al. (2021), our study found no predictors from syntactic complexity indices. This insignificant result could be explained by learners' concentrated overall proficiency. As L2 learners grow more proficient, they generate more complex syntactic structures, but syntactic complexity would become less indicative of proficiency change, especially for intermediate-to-high learners as evidenced by Barrot and Agdeppa (2021).

6. Conclusion

The current study examined cross-genre differences in lexical and syntactic complexity, as well as the generalizability of predictors for writing scores across different task types. The findings substantiate an interaction between task types and linguistic complexity performance. Specifically, more sophisticated and diverse vocabulary was observed in argumentative essays, while descriptive essays contained less complexity in this regard. Additionally, the results indicated that students tended to use more complex syntactic structures in argumentative and descriptive tasks compared to interactive and narrative tasks. Only lexical complexity indices were found to be indicative of overall human ratings, with the strongest predictor of writing scores across all four task types being the Lexical Density Type-Token Ratio (LDTRc). Moreover, the findings suggest that these results can be generalized to tasks involving different generic processes, namely narration, description, interaction, and argumentation. This study contributes to a deeper understanding of how task types influence linguistic complexity in L2 writing.

The implications of these findings are significant for L2 writing assessments and classroom-based teaching. First, since syntactic complexity was not found to predict writing quality changes across the four task types, it is recommended that students with intermediate-to-high proficiency levels focus on improving their vocabulary usage in L2 writing. Second, the relatively low lexical complexity observed in descriptive essays highlights a potential deficiency in learners' familiarity with descriptive writing and suggests room for improvement in their graph interpretation skills. Third, the finding that students produced less sophisticated vocabulary in narrative essays reveals weaknesses in narrative writing, indicating an area for further development.

However, several limitations of this study should be noted. First, the sample consisted solely of intermediate-to-high proficiency English majors, which may limit the generalizability of the findings to less proficient learners. Second, the study did not include instructive or expository writing tasks, due to the limited presence of instructive writing and the presumed similarity in the generic processes of expository writing to other task types. Third, this study focused exclusively on linguistic complexity, excluding other linguistic aspects such as accuracy, fluency, and cohesion. Future research should explore a wider range of task types and linguistic dimensions, utilizing a larger sample size to enhance the robustness and applicability of the findings.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

ORCID iD 0009-0002-2872-4898

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Barrot, J. S., & Agdeppa, J. Y. (2021). Complexity, accuracy, and fluency as indices of college-level L2 writers' proficiency. *Assessing Writing*, 47, 100510. <https://doi.org/10.1016/j.asw.2020.100510>
- [2] Fathi, J., & Rahimi, M. (2022). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: A case of EFL students. *Computer Assisted Language Learning*, 35(7), 1668-1706. <https://doi.org/10.1080/09588221.2020.1825097>
- [3] Fitzgerald, J., & Spiegel, D. L. (1986). Textual cohesion and coherence in children's writing. *Research in the Teaching of English*, 263-280. <https://www.jstor.org/stable/40171083>
- [4] Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5), 223-234. <https://doi.org/10.3102/0013189X11413260>
- [5] Halliday, M. A. K. & Hasan, R. (1989). *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective* (2nd Ed.). Oxford: Oxford University Press. <https://doi.org/10.2307/3586740>
- [6] Halliday, M. A. K., & Matthiessen, C. M. I. M. (2013). *Halliday's Introduction to Functional Grammar*. Routledge. <https://doi.org/10.4324/9780203431269>
- [7] Housen, A., Vedder, I., & Kuiken, F. (2012). *Dimensions of L2 performance and proficiency*. Amsterdam: John Benjamins Publishing Company. <http://digital.casalini.it/9789027273260>
- [8] Jeong, H. (2017). Narrative and expository genre effects on students, raters, and performance criteria. *Assessing writing*, 31, 113-125. <https://doi.org/10.1016/j.asw.2016.08.006>

- [9] 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>
- [10] Khushik, G. A., & Huhta, A. (2020). Investigating syntactic complexity in EFL learners' writing across common European framework of reference levels A1, A2, and B1. *Applied Linguistics*, 41(4), 506-532. <https://doi.org/10.1093/applin/amy064>
- [11] Kuiken, F. (2023). Linguistic complexity in second language acquisition. *Linguistics Vanguard*, 9(s1), 83-93. <https://doi.org/10.1515/lingvan-2021-0112>
- [12] Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, 34, 12-24. <https://doi.org/10.1016/j.jslw.2016.10.003>
- [13] Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied linguistics*, 35(5), 607-614. <https://doi.org/10.1093/applin/amu04>
- [14] Latifi, S., & Gierl, M. (2021). Automated scoring of junior and senior high essays using Coh-Metrix features: Implications for large-scale language testing. *Language Testing*, 38(1), 62-85. <https://doi.org/10.1177/0265532220929918>
- [15] Lee, C., Ge, H., & Chung, E. (2021). What linguistic features distinguish and predict L2 writing quality? A study of examination scripts written by adolescent Chinese learners of English in Hong Kong. *System*, 97, 102461. <https://doi.org/10.1016/j.system.2021.102461>
- [16] 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>
- [17] Miao, Y. (2006). Validating a simulated test of CET 4. *The Asian EFL Journal*, 20(4), 409-439.
- [18] Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics*, 24(4), 492-518. <https://doi.org/10.1093/applin/24.4.492>
- [19] Peng, Y., Sun, J., Quan, J., Wang, Y., Lv, C., & Zhang, H. (2023). Predicting Chinese EFL Learners' Human-rated Writing Quality in Argumentative Writing Through Multidimensional Computational Indices of Lexical Complexity. *Assessing Writing*, 56, 100722. <https://doi.org/10.1016/j.asw.2023.100722>
- [20] Plakans, L., Gebriel, A., & Bilki, Z. (2019). Shaping a score: Complexity, accuracy, and fluency in integrated writing performances. *Language Testing*, 36(2), 161-179. <https://doi.org/10.1177/0265532216669537>
- [21] Polat, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied linguistics*, 35(2), 184-207. <https://doi.org/10.1093/applin/amt013>
- [22] Pu, L., Heng, R., & Cao, C. (2022). The effects of genre on the syntactic complexity of argumentative and expository writing by Chinese EFL learners. *Frontiers in Psychology*, 13, 1047117. <https://doi.org/10.3389/fpsyg.2022.1047117>
- [23] Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research*, 27(3), 343-360. <https://doi.org/10.1177/0267658310395851>
- [24] Shin, S. Y., & Ewert, D. (2015). What accounts for integrated reading-to-write task scores?. *Language Testing*, 32(2), 259-281. <https://doi.org/10.1177/0265532214560257>
- [25] Tabari, M. A., Lu, X., & Wang, Y. (2023). The effects of task complexity on lexical complexity in L2 writing: An exploratory study. *System*, 114, 103021. <https://doi.org/10.1016/j.system.2023.103021>
- [26] Tabari, M. A., & Johnson, M. D. (2023). Exploring new insights into the role of cohesive devices in written academic genres. *Assessing Writing*, 57, 100749. <https://doi.org/10.1016/j.asw.2023.100749>
- [27] Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.
- [28] Yang, H. C. (2016). Describing and interpreting graphs: The relationships between undergraduate writer characteristics and academic graph writing performance. *Assessing writing*, 28, 28-42. <https://doi.org/10.1016/j.asw.2016.02.002>
- [29] Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, 28, 53-67. <https://doi.org/10.1016/j.jslw.2015.02.002>
- [30] 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>
- [31] Zhang, X., & Li, W. (2023). Remodeling effects of linguistic features on L2 writing quality of two genres. *Reading and Writing*, 37(5), 1209-1234. <https://doi.org/10.1007/s11145-023-10438-y>
- [32] Zheng, Y., & Barrot, J. S. (2024). Syntactic complexity in second language (L2) writing: Comparing students' narrative and argumentative essays. *System*, 123, 103342. <https://doi.org/10.1016/j.system.2024.103342>
- [33] Zheng, Y., & Cheng, L. (2008). Test review: college English test (CET) in China. *Language Testing*, 25(3), 408-417. <https://doi.org/10.1177/026553220809243>