

# What Linguistic Features Predict High-Quality Writing

Mingbin Wang

Jiangsu Second Normal University, Nanjing 211200, Jiangsu, China

*“Good writing is clear thinking made visible.”  
- William Wheeler*

WRITING is a basic component of language proficiency, acting as a vital medium for the representation of one’s thoughts, perspectives, attitudes, and more. English writing is a compulsory skill in learning and using the language, and writing proficiency is indicative of the holistic linguistic ability of the individual. Nevertheless, writing instruction has been widely perceived as a major challenge in English language education. Therefore, encouraging debates on “what makes a good English essay” is potentially beneficial for teachers and learners successfully addressing this challenge.

Researchers worldwide have undertaken extensive explorations of linguistic features affecting writing quality from multiple perspectives, including vocabulary, syntax, and discourse. At the lexical level, lexical sophistication is usually deemed a positive predictor of writing quality, with higher-quality essays often containing more sophisticated words that are acquired in higher grades, low-frequency words, and academic vocabulary (Crossley & Kim, 2022). At the syntactic level, there is a positive correlation between syntactic complexity and writing quality. However, many studies find that traditional metrics of syntactic complexity are insufficiently effective. There have been two traditional approaches to its evaluation: (i) assessing syntactic complexity by measuring the sentence length, with the use of longer sentences being recognized as an indicator of syntactic maturity; and (ii) calculating the number of T-units, where a T-unit is defined as one main clause plus all subordinate clauses attached to it, to measure the complexity of subordination. Certain researchers contended that, compared to clausal complexity, phrasal complexity, particularly noun phrase modification (by attributive adjectives and post-noun-modifying prepositional phrases) tends to contribute more to essay quality (Taguchi et al., 2013). In the dimension of discourse cohesion and coherence, a significant positive correlation between the use of global coherence devices and writing quality has been established (Crossley, 2020; Crossley & Kim, 2022). Higher-quality writing typically exhibits

---

© 2026 Insights Publisher. All rights reserved.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License

(<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed by the Insights Publisher.

higher levels of cohesion and coherence, with higher lexical and semantic overlap across paragraphs, thereby foregrounding the essay's theme. Crossley's (2020) research findings, based on an extensive review of relevant empirical studies, confirm the correlations between the three linguistic features (lexical sophistication, syntactic complexity, and text coherence) and the quality of English writing and reveal that the effects of these features are contingent on the essay's topic, genre, and the student's discipline. Wang et al.'s study (2024) further revealed correlations between individual linguistic features and the quality of essays in different genres: significant positive correlations between syntactic complexity and the quality of argumentative essays, between syntactic accuracy and the quality of narrative essays, and between lexical sophistication and the quality of writing in the two said genres.

When looking at the question from the perspective of "evaluation actors," researchers discovered a substantial discrepancy between students' and teachers' perceptions of the criteria for a "well-written essay." Specifically, students tend to focus more on overt features, such as lexical sophistication and syntactic complexity, while teachers place greater emphasis on implicit features like logical coherence, depth of argumentation, and discourse organization. For instance, Ding & Zhao's (2019) study finds that teachers put significantly higher value on essay structure and content than students do, whereas students prioritize the range of vocabulary in writing. The difference in their views leads to the divide in essay evaluation standards between them. It is noteworthy that students' perceptions of writing quality could be unfavorable for their developing effective English writing techniques. Therefore, it is necessary for teachers to make explicit the evaluation criteria and factors influencing essay quality in the teaching process to direct students towards legitimate efforts.

Discourse Cohesion and Coherence and the Writing Quality of English Argumentative Essays: An Analysis Based on Coh-Metrix in this issue investigates the relationships between discourse cohesion and coherence and essay quality. It leveraged Coh-Metrix to quantitatively analyze textual features of 386 English argumentative essays using 27 indicators in four dimensions: connectives, referential cohesion, latent semantic analysis, and situation models; The writing quality of these essays was evaluated by both human raters and the JuKu Grading platform. The results of correlation analysis and multiple linear regression analysis reveal that there were statistically significant correlations between a portion of the 27 indicators and the quality of the essays sampled and that these indicators had varied predictive effects on writing scores (Xie, 2026). Despite its constrained generalizability due to its sample being confined to English essays by Chinese EFL learners, the study provides valuable empirical evidence for improving English writing education, particularly training in use of connectives, referential cohesion, and semantic coherence.

## References

- Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3), 415-443. DOI: <https://doi.org/10.17239/jowr-2020.11.03.01>
- Crossley, S. A., & Kim, M. (2022). Linguistic features of writing quality and development: A longitudinal approach. *The Journal of Writing Analytics*, 6(1), 59-93. DOI: <https://doi.org/10.37514/JWA-J.2022.6.1.04>
- Ding, Y. & Zhao, T. (2019). Chinese University EFL Teachers' and Students' Beliefs About EFL Writing: Differences, Influences, and Pedagogical Implications. *Chinese Journal of Applied Linguistics*, 42(02), 163-181+263. DOI: <https://doi.org/10.1515/CJAL-2019-0010>
- Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *Tesol Quarterly*, 47(2), 420-430. DOI: <https://doi.org/10.1002/tesq.91>

- Wang, J., Kim, Y. S. G., & Cho, M. (2024). Linguistic features in narrative and opinion genres and their relations to writing quality in fourth grade writing. *Journal of Research in Reading*, 47(2), 220-239. DOI: <https://doi.org/10.1111/1467-9817.12453>
- Xie, Y. (2026). Discourse cohesion and coherence and the writing quality of English argumentative essays: An analysis based on coh-metrix. *Science Insights Education Frontiers*, 32(2), 5249-5264. DOI: <https://doi.org/10.15354/sief.26.or178>

**Correspondence to:**

Mingbin Wang  
Jiangsu Second Normal University  
Nanjing 211200, Jiangsu  
China  
E-mail: [396517407@qq.com](mailto:396517407@qq.com)

**Conflict of Interests:** None

**Doi:** 10.15354/sief.26.co014