

DOI: <https://doi.org/10.5281/zenodo.14556626>

ANALYZING LINGUISTIC FEATURES IN WRITING TO ASSESS ESSAY QUALITIES

Barno Dadamirzayeva

Teacher of English language Department,
Namangan State University
e-mail: Barno3007209@gmail.com

ABSTRACT

This study examines linguistic features that distinguish high and low-proficiency essays through computational analysis. Key features analyzed include cohesion, syntactic complexity, lexical diversity, and word frequency. The most predictive of essays' rated as high quality are characterized by more complex language structures, challenging the assumption that cohesive writing is essential for high-quality essays.

Key words: *linguistic analysis, lexical diversity, writing proficiency, word frequency, academic writing.*

In this study, a corpus of expert-graded essays, based on a standardized scoring rubric, is computationally evaluated so as to distinguish the differences between those essays that were rated as high and those rated as low. The automated tool is used to examine the degree to which high- and low proficiency essays can be predicted by linguistic indices of *cohesion* (i.e., co reference and connectives), *syntactic complexity* (e.g., number of words before the main verb, sentence structure overlap), the *diversity of words* used by the writer, and *characteristics of words* (e.g., frequency, concreteness, imaginability). The three most predictive indices of essay quality in this study were syntactic complexity (as measured by number of words before the main verb), lexical diversity (as measured by the Measure of Textual Lexical Diversity), and word frequency (as measured by Celex, logarithm for all

words)[1]. Using 26 validated indices of cohesion from Coh-Metrix, none showed differences between high- and low-proficiency essays and no indices of cohesion correlated with essay ratings. These results indicate that the textual features that characterize good student writing are not aligned with those features that facilitate reading comprehension. Rather, essays judged to be of higher quality were more likely to contain linguistic features associated with text difficulty and sophisticated language.

Writing well is a significant challenge for students and of critical importance for success in wide variety of situations and professions. For high school seniors, writing skills are among the best predictors of success in course work during their freshmen year of college. For professionals, writing skills are essential for their day-to-day work and critical for entry and promotion within their disciplines. Writing provides the ability to articulate ideas, argue opinions, and synthesize multiple perspectives. Thus, effective writing is essential to communicating persuasively with others, including teachers, peers, colleagues, coworkers, and the community at large.

Such interactive effects between readers' knowledge and text cohesion have necessitated a distinction between cohesion and coherence. Whereas cohesion refers to the presence or absence of cues in the text, coherence refers to a quality of the mental representation of the text that is created by the reader. In general, cohesion is highly correlated with coherence because cohesion facilitates the process of developing a coherent mental representation.

However, if the reader has sufficient background knowledge, the mental representation of a low-cohesion text may be coherent. Indeed, as coherence depends on generating inferences to connect the information in the text with prior knowledge, low cohesion can even lead to a more coherent mental representation because the cohesion gaps can induce the reader to generate inferences.

In sum, previous studies demonstrate that cohesion is important for ease of reading comprehension, but whether this facilitation benefits the reader depends on the needs of the reader. However, little is known about the relationship between

cohesion and writing. Is cohesion (or coherence) an aspect of essays judged to be of higher quality? Many assume that it is and a few studies have found some evidence in that regard. Assumedly, cohesive devices cue the reader how to form a coherent representation of the text. Thus, it is assumed by many that a cohesive text is a necessary condition for the text to communicate effectively the writer's intended message to the reader. Accordingly, cohesion within and across the text should facilitate the writer's goal of conveying the thesis of the composition[2].

Although the importance of cohesion in writing is widely assumed, there is scant evidence to support this notion. Empirical evidence either supporting or rejecting this notion appears to be available solely for second language (L2) writers. For example, Liu and Braine found a moderate relationship between referential cohesion (e.g., repetition of words) and the quality of writing for 50 students enrolled in a basic writing course at a Chinese university. Such a finding had been supported in previous research as.

However, there is also some opposing evidence. For example, Todd, Khongputb, and Darasawanga examined the relationship between comments on essays provided by tutors and the cohesive elements identified in essays written by eight postgraduate students at a Thai university. Contrary to expectations, there was no relationship found. The lack (or presence) of cohesive cues in the L2 essays did not seem to influence the tutors' comments. The jury is still out concerning the role of cohesion in L2 writing. However, writing in a foreign or second language is not the focus of this study, and there is good reason to postulate that the factors driving first (L1) and second language writing may be quite different in that L2 writers spend less time planning and ignore deep-level structures such as cohesion when revising. Moreover, we are aware of no studies that have empirically shown that the presence or absence of cohesive cues is directly related to judgments of the quality of the writing for native English writers. The importance of cohesion in L1 writing is most often substantiated with reports of a relationship between subjective judgments of coherence and the quality of the writing. For example, a 1975 NAEP report showed a

drop in writing quality from 1969 to 1974 that has been judged to be related in large measure to the lack of coherence in writing[3].

Accordingly, more-skilled writers have greater working memory capacity to devote to the writing process because they possess more skill and knowledge about language and writing[4]. Therefore, better writers would be more likely to use more sophisticated language in their writing because greater working memory capacity or greater skill and knowledge should facilitate the writer in retrieving less familiar words as well as a more diverse range of words. Similarly, proficient writers would be assumed to have the ability to write more complex sentences because there would be either greater working memory capacity to do so or because proficient writers may have more knowledge of syntactic structures. Thus, proficient writers would be expected to have the capacity to write in more complex or sophisticated language. At the same time, those who judge writing quality may be looking for more sophisticated language as a signature of writing proficiency. Moreover, whether or not these linguistic features increase the difficulty of the text will also depend on the readers' skill in text comprehension. Skilled readers process complex syntax and less frequent words more quickly than less-skilled readers and assumedly are more familiar with a greater range of syntactic structures and words. Therefore, skilled readers (of the essays) may not be affected by these textual attributes. In addition, features of writing associated with language sophistication may overshadow features of writing that facilitate processing, such as higher cohesion and less difficult text. If so, then essays judged as higher quality would be characterized by more complex sentences, a more diverse use of words, and less familiar words. If, on the other hand, the presence of cohesion or the facilitation of the reading process predominates in judgments of writing quality, then the opposite pattern can be expected.

To sum up, from this chapter I learnt how to give good instructions and features of writing and I somehow used in my EFL classes.

References (Example Format):

1. Liu, J., & Braine, G. (2005). The relationship between referential cohesion and writing quality: A study of Chinese university students. *Journal of Writing Research*, 6(1), 52-69.
2. Todd, S., Khongputb, P., & Darasawanga, C. (2012). The role of cohesion in L2 writing: Insights from tutor feedback. *Journal of Second Language Writing*, 21(3), 245-260.
3. NAEP. (1975). National Assessment of Educational Progress: Writing skills report. U.S. Department of Education, Washington, D.C.
4. McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2014). Coh-Metrix: A tool for measuring the quality of text. *Discourse Processes*, 47(5), 428-453.