Instituto Internacional de Investigación y Desarrollo Tecnológico Educativo INDTEC, C.A.
DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

Lexical Richness and Academic Writing Quality in English Compositions

Ángeles María Zambrano Amores[1]; Ricardo Esteban Llongo Cruz[1]; Rodrigo Vicente Tovar Viera[1]
[1]Universidad Técnica de Cotopaxi (UTC), Latacunga, Ecuador
angeles.zambrano2230@utc.edu.ec | https://orcid.org/0009-0007-9783-2450
ricardo.llongo9352@utc.edu.ec | https://orcid.org/0009-0004-8860-0074
rodrigo.tovar@utc.edu.ec | https://orcid.org/0000-0002-5960-3020
Correspondence: angeles.zambrano2230@utc.edu.ec

Abstract: Lexical richness is considered an indicator of second language writing development, although its contribution to the quality of academic writing in English as a foreign language (EFL) remains unclear. This study aimed to determine the direction and magnitude of Pearson correlations between lexical diversity, density, and sophistication and four components of academic writing quality (coherence, argumentation, clarity, and precision). Using a deductive method, a quantitative approach, and a nonexperimental, cross-sectional, correlational design, the study analyzed the argumentative essays of 24 B1-level students from an Ecuadorian university, selected through convenience sampling. Diversity was measured with the type-token ratio (TTR), density with the proportion of content words, and sophistication with Academic Word List (AWL) coverage; quality was assessed with an analytic rubric scored from 1 to 10 per component by three raters, with no agreement coefficient reported. The coefficients showed negligible or weak associations between diversity and the four components, and weak to moderate associations for density and sophistication; five of the twelve tests were significant ($p < .05$), but none remained so after Bonferroni correction. The study concludes that, in this group, lexical richness is associated to a limited extent and with varying magnitude with academic writing quality; the conjecture that vocabulary operates together with discourse features requires studies that measure them independently.

Keywords: lexical richness; lexical diversity; lexical density; lexical sophistication; academic writing quality.
UNESCO International Classification Code: 5701.11 - Teaching of languages.
OECD-FOS Classification: 5.3 - Educational sciences.

Riqueza Léxica y Calidad de la Escritura Académica en Composiciones en Inglés

Resumen: La riqueza léxica se considera un indicador del desarrollo de la escritura en una segunda lengua, aunque su contribución a la calidad de la escritura académica en inglés como lengua extranjera (EFL) permanece poco precisada. Este estudio tuvo como objetivo determinar el sentido y la magnitud de las correlaciones de Pearson entre la diversidad, la densidad y la sofisticación léxicas y cuatro componentes de la calidad de la escritura académica (coherencia, argumentación, claridad y precisión). Con un método deductivo, un enfoque cuantitativo y un diseño no experimental, transversal y correlacional, se analizaron los ensayos argumentativos de 24 estudiantes de una universidad ecuatoriana, de nivel B1, seleccionados por conveniencia. La diversidad se midió con el índice tipo-token (TTR), la densidad con la proporción de palabras de contenido y la sofisticación con la cobertura de la Academic Word List (AWL); la calidad, con una rúbrica analítica de 1 a 10 por componente aplicada por tres evaluadores, sin coeficiente de concordancia reportado. Los coeficientes mostraron asociaciones despreciables o débiles de la diversidad con los cuatro componentes y de débiles a moderadas de la densidad y la sofisticación, significativas en cinco de los doce contrastes ($p < ,05$), ninguna tras la corrección de Bonferroni. Se concluye que, en este grupo, la riqueza léxica se asocia de manera limitada y con magnitud variable con la calidad de la escritura académica; la conjetura de que el vocabulario opera junto con rasgos discursivos requiere estudios que los midan de forma independiente.

Palabras clave: riqueza léxica; diversidad léxica; densidad léxica; sofisticación léxica; calidad de la escritura académica.
Código de clasificación internacional UNESCO: 5701.11 - Enseñanza de lenguas.
Clasificación OCDE-FOS: 5.3 - Ciencias de la educación.

Riqueza lexical e qualidade da escrita acadêmica em composições em inglês

Resumo: A riqueza lexical é considerada um indicador do desenvolvimento da escrita em segunda língua, embora sua contribuição para a qualidade da escrita acadêmica em inglês como língua estrangeira (EFL) permaneça pouco esclarecida. Este estudo teve como objetivo determinar o sentido e a magnitude das correlações de Pearson entre a diversidade, a densidade e a sofisticação lexicais e quatro componentes da qualidade da escrita acadêmica (coerência, argumentação, clareza e precisão). Com um método dedutivo, uma abordagem quantitativa e um delineamento não experimental, transversal e correlacional, analisaram-se os ensaios argumentativos de 24 estudantes de uma universidade equatoriana, de nível B1, selecionados por conveniência. A diversidade foi medida com a razão tipo-token (TTR), a densidade com a proporção de palavras de conteúdo e a sofisticação com a cobertura da Academic Word List (AWL); a qualidade, com uma rubrica analítica de 1 a 10 por componente aplicada por três avaliadores, sem coeficiente de concordância relatado. Os coeficientes mostraram associações desprezíveis ou fracas da diversidade com os quatro componentes e de fracas a moderadas da densidade e da sofisticação, significativas em cinco dos doze testes ($p < 0,05$), nenhuma após a correção de Bonferroni. Conclui-se que, neste grupo, a riqueza lexical se associa de maneira limitada e com magnitude variável com a qualidade da escrita acadêmica; a conjectura de que o vocabulário opera em conjunto com traços discursivos requer estudos que os mensurem de forma independente.

Palavras-chave: riqueza lexical; diversidade lexical; densidade lexical; sofisticação lexical; qualidade da escrita acadêmica.
Código de Classificação Internacional da UNESCO: 5701.11 - Ensino de línguas.
Classificação OCDE-FOS: 5.3 - Ciências da educação.

How to cite this article:
Zambrano, Á. M., Llongo, R. E., & Tovar, R. V. (2026). Lexical Richness and Academic Writing Quality in English Compositions. *Revista Scientific, 11*(41), 181–196. https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196

Date Received:	Date Accepted:	Date Published:
03-12-2025	10-02-2026	05-08-2026

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

1. Introduction

Academic writing requires learners of English as a foreign language (EFL), especially in higher education, to produce coherent, clear, and well-argued texts using vocabulary appropriate to the academic register. Writing an argumentative essay involves selecting words that sustain a position, connect ideas, and convey meanings precisely, so the quality of a text may depend, among other factors, on the lexical decisions of the writer (Yang et al., 2023).

University learners may face specific difficulties with academic vocabulary. A case study of five advanced-level Chinese EFL students documented limited knowledge of the academic register, a small academic vocabulary size, and problems with word choice and collocations when writing argumentative essays (Min & Sukying, 2024). Features such as lexical sophistication and syntactic complexity have also been analyzed in the academic writing of adolescent second language learners (Maamuujav et al., 2021).

One line of research on second language vocabulary has focused on receptive vocabulary knowledge by frequency level, assessed with instruments such as the updated version of the Vocabulary Levels Test (Webb et al., 2017). These measures indicate which frequency levels a learner knows, but they do not describe how the learner uses words when producing an extended text. In view of that limitation, attention has broadened toward lexical richness, which characterizes the vocabulary actually used in written production (Laufer & Nation, 1995; Read, 2000).

Despite sustained interest, the contribution of lexical richness to the quality of academic writing in EFL remains inconclusive. Some studies report that certain lexical indices explain an appreciable share of the variation in essay scores (Yang et al., 2023), but the results differ depending on the index and the context. Likewise, several recent studies explain the variation in an overall writing score, which makes it impossible to know which of its components each lexical feature is related to (Susoy & Durmuşoğlu, 2024; Yang et al., 2023).

1.1. Lexical Richness as a Multidimensional Construct

Lexical richness refers to the vocabulary that is effectively deployed in a text and constitutes a relevant dimension for assessing the competence of second language learners (Laufer & Nation, 1995; Read, 2000). Building on that definition, lexical production is examined through four dimensions: vocabulary diversity, the presence of advanced words related to the topic or context, the proportion of content words, and the number of lexical errors (Hao et al., 2023). This study takes up the first three.

Lexical diversity refers to the degree to which a writer draws on a varied repertoire of words and is usually estimated as the ratio of distinct words (types) to total words (tokens), known as the type-token ratio (TTR). Its main limitation is its sensitivity to text length, because values vary with the amount of data (Raufi et al., 2024): as a text grows longer, words are repeated and the ratio declines. To mitigate this problem, corrected versions have been proposed, such as the corrected TTR, which some studies include (Abu-Rabiah, 2023; Yang & Zheng, 2024). In this study, the uncorrected TTR was used.

Lexical density expresses the proportion of content words, that is, nouns, lexical verbs, adjectives, and adverbs, relative to the total number of words in the text, which is the operational definition adopted here. In systemic functional grammar, density is usually calculated per clause, and written discourse is characterized by high density because its lexical words concentrate information, which links density to the formality and informational orientation of the academic register (Halliday & Matthiessen, 2013). In comparisons with native writers, Chinese university EFL students reach density levels comparable to those of native writers (Yang & He, 2025).

Lexical sophistication denotes the use of advanced vocabulary, whether because of its low frequency or its membership in a specialized register. A

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

common way of measuring it is to calculate coverage of the Academic Word List (AWL), an inventory of word families that are frequent in academic texts and that excludes those with the highest general frequency (Coxhead, 2000). Another is to build frequency profiles, such as the Lexical Frequency Profile (Laufer & Nation, 1995); both procedures are automated with VocabProfile (Cobb, 2007). Here, sophistication is measured only with the AWL, which captures general academic vocabulary rather than the rarity of words.

Beyond the AWL, recent research has expanded the repertoire of sophistication indicators. Indices sensitive to word meaning (Lu & Hu, 2022), second language and first language semantic context indices (Monteiro et al., 2023), and multidimensional models in which frequency, age of acquisition, and naming reaction time explain the difficulty of academic words (Vitta et al., 2023) have been proposed. It follows that sophistication cannot be reduced to a single property of words.

In line with Zheng (2025), lexical richness is described through three distinct dimensions, diversity, density, and sophistication, and in a corpus of Japanese political speeches no consistent relationship was observed among them, nor was an effect of text length on density and sophistication. Although that corpus differs from essays by learners of English, the finding is a reminder that the three dimensions are not interchangeable and that each requires its own interpretation.

1.2. Academic Writing Quality and Its Assessment

Academic writing quality is a multidimensional construct that is not exhausted by grammatical accuracy. Writing assessment models distinguish, among other aspects, content and idea development, organization and coherence, and vocabulary and language control, and they recommend analytic scales, which score each dimension separately, when diagnostic information about writers' performance is needed (Weigle, 2002). This type of scale makes it possible to relate each linguistic feature to a specific component of quality.

This study considers four components of quality. Coherence is understood as the logical sequencing and organization of ideas; argumentation refers to the development and support of claims; clarity relates to the comprehensibility of the text; and precision refers to grammatical and lexical appropriateness. These components are interdependent, and analyzing them separately is an analytic decision whose empirical distinctness was not tested. Because precision includes lexical appropriateness, its association with the lexical indices may partly reflect an overlap between criterion and predictor.

The assessment of the components depends on the judgment of human raters, who do not always prioritize the same features. Following Susoy and Durmuşoğlu (2024), who interviewed eight teachers with extensive experience in scoring essays by prospective English teachers, some raters showed a nuanced understanding of syntactic complexity and lexical diversity, whereas others gave more weight to content and overall organization. This variability justifies the use of rubrics that make the criteria for each component explicit.

Complementarily, automated lexical richness indices are used as predictors of the scores assigned by human raters (Yang et al., 2023). For their part, Appel and Arias (2023), working with integrated essays from an academic English test, found that total text length, an indicator of writing fluency, stood out among the predictors, and the authors point to the need for a broader range of metrics to account for writing proficiency. This observation cautions against attributing the explanatory weight of quality to a single lexical dimension.

1.3. Empirical Evidence on the Relationship Between Both Constructs

The evidence on the relationship between lexical richness and writing quality is heterogeneous. In expository essays by Chinese EFL students, the four groups of lexical measures, density, sophistication, variation, and fluency, correlated significantly with

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

scores, albeit with low to medium magnitudes, and two indices, word count and noun variation, explained 38.5% of the variance (Yang et al., 2023). Results thus vary according to the index and the type of text.

Other studies agree that sophistication is linked to quality. In analytical texts by adolescent second language learners, syntactic complexity and lexical sophistication predicted quality scores (Maamuujav et al., 2021), and in expository essays by EFL university students, the frequency and range of content words distinguished proficiency levels, unlike lexical diversity (Ait Hammou et al., 2023). The two studies come from different contexts and used different indices, and for that reason their results are not directly comparable.

These findings are qualified when other levels and contexts are examined. In 309 argumentative essays by Turkish students in English teacher education, diversity, density, and vocabulary size were analyzed, and only lexical diversity correlated significantly with first-year scores, with an explained variance of 7.8%, and with no significant effects on fourth-year scores; the authors conclude that lexical features are important but are not the only determinants of scores (Tömen & Durmuşoğlu, 2023).

A pattern opposite to what was expected appeared in a study with Indonesian university students in online academic writing, whose abstract reports that lexical diversity, expressed through the D index, correlated negatively and significantly with the use of academic vocabulary (Raufi et al., 2024). The statistics in the results section do not clearly support that correlation, and the finding is taken with caution. Because that study related two lexical measures and not quality scores, it provides no evidence on writing quality.

The comparison across contexts also yields differences. Compositions by secondary school students in Hong Kong, who learn English as a second language, and by students from mainland China, who learn it as a foreign language, differed in their lexical and syntactic features, with greater lexical diversity in the second group and greater content density in the first (Zhang & Kang, 2022). In turn, Chinese university EFL students show lower lexical sophistication and variation than native English writers (Yang & He, 2025), which places the question in the realm of learning conditions.

Research on lexical development offers another line of analysis. In a corpus of more than 1,900 essays by Polish adolescent EFL learners, lexical variation developed faster than density and sophistication (Rokoszewska, 2024), and in a one-year study of Arabic-Hebrew bilinguals who wrote argumentative essays in their second language, no significant increase in lexical density or diversity was found, although students with lower diversity improved notably at the second measurement (Abu-Rabiah, 2023).

Taken together, the literature reviewed offers mixed evidence: several studies link sophistication with quality, whereas the results on lexical diversity are inconsistent and depend on the context, text length, and the index used. Along these lines, a model aimed at reducing redundancy among indices retained six, among them lexical density, lexical sophistication, and the corrected TTR, as the most sensitive for distinguishing levels among Chinese university EFL students (Yang & Zheng, 2024).

The applicability of the indices has also been examined across languages. Of 44 indices, most developed for Indo-European languages, five, among them the number of distinct words and the Lexical Frequency Profile, effectively predicted the writing quality of learners of Chinese as a foreign language (Hao et al., 2023). In another study, Hu et al. (2022), using narratives by Korean learners of Chinese, found that several phraseological measures correlated more strongly with quality than those based on topic and comment. Both studies address Chinese and other indices, and they only suggest that quality is linked to several indices rather than to a single one.

The mixed evidence described leads to a conjecture that guides the reading of the results: vocabulary might operate together with discourse features, such as cohesion between sentences and the

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

rhetorical structure of the essay, rather than as an isolated correlate of quality. The coherence and argumentation scores of the rubric are expert judgments of the whole text and not independent measurements of those features, so the study can neither test the conjecture nor assume that vocabulary alone explains the quality of the essays. Within this framework, the study examines which dimension of lexical richness is associated with each component of quality.

1.4. Research Problem, Question, and Objective

Among EFL students at an intermediate level who are training to become English teachers, there remains uncertainty about the extent to which lexical diversity, density, and sophistication are associated with specific components of quality, such as coherence, argumentation, clarity, and precision, in short argumentative essays. Describing those associations would make it possible to formulate conjectures about which vocabulary features might be related to text quality and to avoid exclusive attention to indicators that could be of little informational value.

The study is justified from three angles. Theoretically, it contrasts three dimensions of lexical richness with four quality components assessed separately, rather than with an overall score. Methodologically, it combines automated lexical measures with an analytic rubric and estimates the precision of the associations through confidence intervals. Pedagogically, it offers descriptive information on which lexical features could be explored in later research on the teaching of writing in English teacher education.

On this basis, the following research question is posed: To what extent is lexical richness associated with the components of coherence, argumentation, clarity, and precision of academic writing in EFL? Given the correlational nature of the problem, the question is framed in terms of association and not of effect, and the answer is expressed through the direction and magnitude of the coefficients and through the precision with which they are estimated. The question thus delimits the scope of the study to the description of associations between lexical dimensions and quality components in a specific group.

To answer it, the general objective of this study is to determine the direction and magnitude of the Pearson correlations between the three dimensions of lexical richness (diversity, density, and sophistication) and the four quality components (coherence, argumentation, clarity, and precision) of argumentative essays written by 24 students at an Ecuadorian university, who are EFL learners at the B1 level, and to identify which of these correlations reach statistical significance. Complementarily, the precision of each coefficient is estimated through its confidence interval, and its robustness to the multiple tests is assessed.

2. Methodology

The methodology was designed to estimate the association between three dimensions of lexical richness and four components of writing quality in the essays of a group of students, with a single text per participant. To this end, lexical measures obtained with an analysis tool and scores assigned with an analytic rubric were combined, and correlation coefficients were calculated together with confidence intervals and a correction for multiple tests. Each decision is justified below, along with the limits that condition its interpretation.

2.1. Method, Approach, and Design

The study followed a deductive method, since it started from established theoretical constructs (lexical richness and writing quality) to examine their association in empirical data (Hernández-Sampieri & Mendoza, 2018). A quantitative approach was adopted, which measures variables with numerical instruments and analyzes their relationships through statistical procedures (Creswell & Creswell, 2017). The design was nonexperimental, cross-sectional, and correlational, because the variables were observed without manipulation, in a single writing task, with the purpose of estimating the direction and magnitude of their

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

association (Cohen et al., 2017).

The three decisions were articulated with one another. The deductive method guided the selection of lexical dimensions and quality components on the basis of the literature; the quantitative approach required operationalizing each variable through an index or a score; and the correlational design determined the scope of the inferences, which are limited to describing associations and do not allow causal relationships or predictions in the strict sense to be established. For that reason, the results are framed in terms of association and estimation precision, and not of effect or hypothesis testing.

The study did not formulate hypotheses. Its scope was exploratory within the correlational framework: the aim was to estimate the direction, magnitude, and significance of twelve associations between lexical dimensions and quality components, without setting an expected value for each. This decision responds to the fact that the previous literature offers inconclusive results and that the small sample size discourages testing specific directional hypotheses, hence the exploratory nature of the findings.

2.2. Context and Participants

The research was conducted at the Universidad Técnica de Cotopaxi (UTC), located in Latacunga, Ecuador. Participants were 24 undergraduate students enrolled in the seventh semester of the Pedagogía de los Idiomas Nacionales y Extranjeros (Pedagogy of National and Foreign Languages) program. All had a B1 level of English, based on the university's institutional placement criteria, and the results are interpreted only for this group of intermediate-level students training to become English teachers.

The selection criterion was nonprobability convenience sampling, determined by the accessibility of the participants and their relevance to the objectives of the study. This criterion is distinct from the characteristics of the sample, which were the semester attended, the program, and the B1 level. Because the sample was a small convenience sample, the results describe the association between the variables in this group and are not extrapolated to the population of English students at the university or in the country.

The authors belong to the university where the study was conducted, and two of them are eighth-semester students in the same program as the participants, one semester ahead of them (seventh). They therefore occupy the position of internal researchers, which may have influenced the participants' willingness and is acknowledged as a limitation. The essays were scored by three expert raters, and the article does not report whether they knew the identity of the participants.

2.3. Variables and Instruments

Data were collected through an academic writing task and a lexical analysis tool. Each student wrote an argumentative essay with a target length of between 200 and 250 words. The task was designed to obtain a sample of the participants' lexical and persuasive skills in an academic writing situation. The short length of the essay conditioned the interpretation of indices that are sensitive to text length, such as the TTR, and for that reason this characteristic is taken up again when the results are assessed.

There were seven variables: three of lexical richness and four of writing quality. Lexical diversity was measured with the type-token ratio (TTR), calculated as the quotient between the number of distinct words and the total number of words in the essay. Lexical density was calculated as the ratio of content words to total words. Lexical sophistication was identified from the proportion of Academic Word List (AWL) words present in each essay, that is, its coverage.

The TTR and AWL coverage were obtained with the VocabProfile tool, which classifies vocabulary by frequency levels and counts types and tokens; this kind of quantification of lexical coverage through computational text analysis is addressed in the literature on second language reading (Cobb, 2007). Lexical

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

density was calculated as the proportion of content words over the total number of words in the essay. The indices were expressed as proportions between 0 and 1, so that higher values indicated greater diversity, a greater proportion of content words, and greater coverage of academic vocabulary.

Writing quality was assessed with an analytic rubric adapted from Weigle (2002), which evaluated the four components defined in the Introduction: coherence, argumentation, clarity, and precision. Each component was scored on a 10-point scale, on which 1 corresponded to a low level and 10 to an excellent level, and was analyzed separately, which made it possible to examine its association with each lexical index. The article does not report the descriptors of each level or the adaptation procedure, and the observed scores ranged between 6 and 9, so they are read as relative positions within the group and not as absolute performance levels.

Scoring was carried out by three raters with expertise in the subject matter, who evaluated each essay on the four components with the analytic rubric, and the scores for each component were integrated into a single value per essay. The study did not include an agreement coefficient among the raters, so the reliability of the rubric could not be assessed, which is a limitation of the study. For that reason, the scores were interpreted as expert judgments of an ordinal nature, and the associations with the lexical indices were read with caution.

2.4. Procedure

The procedure followed a structured sequence. First, after giving verbal informed consent, the students wrote the argumentative essay corresponding to the writing task. Second, all essays were manually cleaned and processed with VocabProfile to extract the TTR and AWL coverage, and the proportion of content words was calculated. Third, the three expert raters evaluated each essay on the four quality components using the analytic rubric, and the scores were organized together with the indices in a single database.

The article does not report the data collection period, the time allowed, or the conditions under which the essays were written, such as permitted resources and supervision, nor the actual length of each text, the version of VocabProfile used, or the criteria for manual cleaning. These omissions delimit the scope of the analysis, since the comparability of length-sensitive indices and the reproducibility of the lexical processing depend on such data, and cleaning criteria can alter the counts, for example, if spelling errors are counted as distinct types.

2.5. Data Analysis

The analysis was carried out in three stages. In the first, the lexical indices and the quality scores were described through the median, quartiles, and extremes, represented in box plots; the values of the lexical indices were obtained by measuring the position of each element of the plots in figure 1 and were rounded to two decimal places, and those of quality were read from figure 2, and the article does not report means or standard deviations. In the second, Pearson correlation coefficients were calculated between each lexical dimension and each quality component, twelve in total. In the third, complementary statistics were derived.

Pearson's coefficient was chosen because of the continuous nature of the lexical indices and the interest in estimating the magnitude and direction of linear associations, although the rubric yields ordinal scores and normality was not checked. According to Schober et al. (2018), this coefficient is used with data that follow a bivariate normal distribution, whereas Spearman's coefficient can be used with ordinal data or with relevant outliers. Because the quality scores take few distinct values, with numerous ties, the p value and the confidence interval are approximate. The significance level was $p < .05$, two-tailed, with 22 degrees of freedom.

The magnitude of the coefficients was interpreted, in line with Schober et al. (2018), with the following scale: negligible (less than .10), weak (.10 to .39), moderate (.40 to .69), strong (.70 to .89), and very strong (.90 to 1.00). The cutoff points of this scale are conventional, and the

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

authors themselves caution that they should be used with judgment; in this study they were used as a rough guide and were complemented with the confidence intervals and the coefficient of determination of each correlation.

From each coefficient and the sample size (n = 24), the following were calculated without the need for additional data: the two-tailed p value, the 95% confidence interval using Fisher's z transformation, with an approximate standard error of $1/\sqrt{21}$, and the coefficient of determination r^2. These values were derived from coefficients rounded to two decimal places; consequently, they are approximate and may differ slightly from those that would be obtained with the individual scores. For that reason, significance is interpreted together with the confidence intervals and the correction for multiple tests, not in isolation.

Because the sample size was determined by the availability of participants, the sensitivity of the test was estimated post hoc, that is, the minimum detectable coefficient with a power of .80, calculated with the exact distribution of the correlation coefficient, with a two-tailed alpha of .05 and with the corrected alpha. Likewise, since twelve coefficients were tested, the critical value of r with the Bonferroni correction (alpha = .05/12) was calculated and compared with the uncorrected critical value. These calculations were used to assess the robustness of the results, and their values are presented in the Results section.

3. Results

The results are presented in two blocks. The first describes the distribution of the three dimensions of lexical richness and the four components of writing quality in the 24 essays analyzed. The second presents the Pearson correlations between the two sets of variables, together with the derived statistics that make it possible to assess their significance, their precision, and their robustness to multiple testing with a small sample. Each block is accompanied by the tables and figures that support the values discussed.

3.1. Description of Lexical Richness and Writing Quality

The descriptive statistics in table 1 summarize the median, quartiles, and extreme values of each lexical index and each quality component. The lexical indices were proportions from 0 to 1, and the quality scores corresponded to the 1 to 10 scale of the analytic rubric applied to each essay, so both groups of variables were interpreted on different scales that are not directly comparable. This description precedes the analysis of the associations and does not by itself imply any relationship between the two groups of variables.

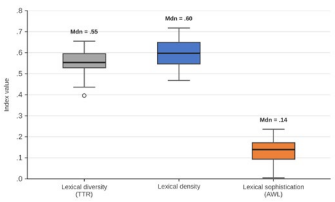
Table 1. Descriptive Statistics of Lexical Richness and Writing Quality.

Variable	Mdn	Q1	Q3	Min	Max
Lexical diversity (TTR)	.55	.53	.60	.40	.66
Lexical density	.60	.55	.65	.47	.72
Lexical sophistication (AWL)	.14	.09	.17	.01	.24
Coherence	8	7	8	6	9
Argumentation	7	6	7.75	6	9
Clarity	8	8	8	6	9
Precision	8	7	8	6	9

Note. Source: Zambrano et al. (2025). *Mdn* = median; *Q1* and *Q3* = first and third quartiles; Min and Max include outliers. The lexical indices are proportions between 0 and 1, and the quality scores are expressed on the rubric scale. n = 24. The values of the lexical indices were obtained by measuring the box plots in figure 1 and were rounded to two decimal places; the quality scores were read from the scale in figure 2, which is integer except for the interpolated quartiles (Q3 of argumentation = 7.75).

Table 1 shows that lexical density had the highest median of the three indices (.60), followed by diversity (.55) and sophistication (.14), whose AWL coverage occupied the lower end of the scale. The interquartile ranges were narrow for diversity (.53 to .60) and for density (.55 to .65). For quality, the medians were 8 for coherence, clarity, and precision, and 7 for argumentation. No score was lower than 6 or higher than 9. The distribution of the three lexical indices is shown in figure 1 by means of box plots, with the median labeled.

Figure 1. Distribution of Lexical Diversity, Density, and Sophistication.



Note. Source: Zambrano et al. (2025). *Mdn* = median. The boxes

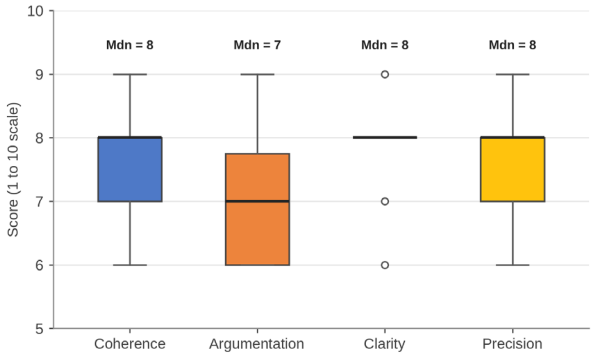
DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

span the first to the third quartile, the whiskers reach the extreme non-outlier values, and the circle marks an outlier. The values of the lexical indices in table 1 were measured on this plot. *n* = 24.

Figure 1 represents that ordering and makes it possible to appreciate three features of the variability that the table summarizes only in part. For diversity, the box was the most compact and showed a single outlier below the lower whisker, which corresponds to the minimum of .40 in table 1. For sophistication, although close to zero, the greatest relative dispersion with respect to its median was observed, with a lower whisker that almost reached zero. For density, the whiskers extended from .47 to .72, an absolute range similar to that of the other two dimensions, and these differed more in their position and relative dispersion than in their absolute range.

The writing quality scores on the four components assessed with the rubric are also represented in figure 2 by means of box plots with the median labeled. Because the scores are represented as integer values on a 10-point scale, several of the lines delimiting the boxes and whiskers coincided at the same value, which explains the appearance of collapsed or truncated boxes in some components and the presence of isolated outliers in clarity.

Figure 2. Distribution of Writing Quality Scores by Component.



Note. Source: Zambrano et al. (2025). *Mdn* = median. Scores are represented as integer values on the 1 to 10 scale, and the circles mark outliers. *n* = 24.

Figure 2 showed that the quality scores were concentrated in a narrow range. Coherence and precision shared the same box, with the median at the upper limit, whereas argumentation showed the lowest median and the box that extended farthest toward the lower values. For clarity, the box was reduced to a line at the median, which indicated that at least half of the texts received the same score, and the cases that departed from it were marked as outliers, both above and below that median, although that marking only indicates that they differ from the common score.

3.2. Associations Between Lexical Richness and Writing Quality

In the correlation stage, with *n* = 24, 22 degrees of freedom, and a two-tailed test at the .05 level, the critical value of *r* was .404, so coefficients whose absolute value equaled or exceeded that value were marked as significant. Magnitude was interpreted with the scale described for the analysis. The coefficients are presented in table 2, and their complementary statistics in table 3. The critical value corresponds to Student's t distribution with those degrees of freedom and was applied equally to all twelve coefficients.

Table 2. Pearson Correlation Coefficients Between Lexical Richness and Writing Quality.

Lexical dimension	Coherence	Argumentation	Clarity	Precision
Lexical diversity (TTR)	.07	.10	.27	.09
Lexical density	.39	.30	.48*	.43*
Lexical sophistication (AWL)	.42*	.44*	.31	.47*

Note. Source: Zambrano et al. (2025). *n* = 24. The asterisk indicates $|r| \geq .404$, the two-tailed critical value with 22 degrees of freedom ($p < .05$). No coefficient reaches .563, the critical value with the Bonferroni correction for twelve tests.

Table 2 shows that lexical diversity yielded coefficients from .07 to .27 with the four components, none of them significant. Lexical density was significantly associated with clarity (*r* = .48) and with precision (*r* = .43), whereas its coefficients with coherence (.39) and argumentation (.30) did not reach the critical value. Lexical sophistication showed significant coefficients with coherence (.42), argumentation (.44), and precision (.47), and a nonsignificant one with clarity (.31). In total, five of the twelve coefficients were significant, and all twelve coefficients were positive.

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

Table 3. p Value, 95% Confidence Interval, and Coefficient of Determination for Each Correlation.

Lexical dimension	Component	*r*	*p*	95% CI	r^2
Lexical diversity (TTR)	Coherence	.07	.745	[−.34, .46]	.005
Lexical diversity (TTR)	Argumentation	.10	.642	[−.32, .48]	.010
Lexical diversity (TTR)	Clarity	.27	.202	[−.15, .61]	.073
Lexical diversity (TTR)	Precision	.09	.676	[−.33, .48]	.008
Lexical density	Coherence	.39	.060	[−.02, .69]	.152
Lexical density	Argumentation	.30	.154	[−.12, .63]	.090
Lexical density	Clarity	.48	.018	[.09, .74]	.230
Lexical density	Precision	.43	.036	[.03, .71]	.185
Lexical sophistication (AWL)	Coherence	.42	.041	[.02, .70]	.176
Lexical sophistication (AWL)	Argumentation	.44	.031	[.04, .72]	.194
Lexical sophistication (AWL)	Clarity	.31	.140	[−.11, .63]	.096
Lexical sophistication (AWL)	Precision	.47	.020	[.08, .73]	.221

Note. Source: Zambrano et al. (2025). CI = confidence interval calculated with Fisher's *z* transformation; r^2 = coefficient of determination. The values were derived from the coefficients in table 2 and from n = 24 (22 degrees of freedom), assume bivariate normality, which was not checked, and were calculated with coefficients rounded to two decimal places, without that rounding changing the classification as significant or nonsignificant.

Table 3 shows that the *p* values of the five significant coefficients ranged between .018 and .041, and that their confidence intervals excluded zero, although with lower limits close to it, between .02 and .09. The intervals of the nonsignificant coefficients included negative and positive values, and that of density with coherence (p = .060) reached up to .69. The coefficients of determination of the significant correlations ranged between .18 and .23, and those of lexical diversity did not exceed .08, which indicated very little shared variance.

The sensitivity and multiple-testing correction calculations yielded the following: with n = 24, the minimum detectable coefficient with a power of .80 was approximately .54, a value higher than the maximum observed (.48). When the Bonferroni correction was applied to the twelve tests, the critical value of *r* rose to .563 (alpha ≈ .0042), and none of the five significant coefficients retained its significance, since the lowest *p* value (.018) did not reach .0042. With that correction, the minimum detectable coefficient with a power of .80 rises to approximately .67, so only associations far larger than those observed would be detected.

4. Discussion

The results indicate that lexical richness is associated to a limited extent and with a magnitude that varies according to the dimension and the component. Five of the twelve coefficients reach significance at the .05 level, all positive and of moderate magnitude, but none retains it after the Bonferroni correction, which, with 24 participants, detects only coefficients of approximately .67 or greater at a power of .80, and the intervals are wide. Consequently, the associations are interpreted as indications compatible with the previous literature and not as conclusive evidence, and the explanations that follow are conjectures that the design does not allow to be tested.

That one association is significant and another is not does not imply that the two differ from each other: with 24 participants the confidence intervals overlap, both across components (the .48 for density with clarity and the .39 with coherence) and across dimensions (the interval for diversity with coherence, [−.34, .46], overlaps with those for density and sophistication). Overlap is only a rough indication, since the twelve correlations come from the same participants. For that reason, the order among the coefficients is taken as descriptive and not as a verified hierarchy.

The coefficients for lexical diversity are the smallest in magnitude, although the intervals do not allow them to be distinguished from those of the other dimensions. This pattern agrees with the finding that diversity did not discriminate between proficiency levels (Ait Hammou et al., 2023), although that study compared levels and did not correlate scores. It differs from what was found among Turkish prospective teachers, where only diversity was associated with first-year scores (Tömen & Durmuşoğlu, 2023), perhaps because of sample size, since with the 309 essays in that study a small association reaches significance more easily than with 24 participants.

Two reasons may explain the low magnitude of the lexical diversity coefficients. The first is the sensitivity of the TTR to text length, which, in essays with a target

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

length of 200 to 250 words and without correction of the index, limits comparability. The second is functional: in a short essay, the repetition of words may serve communicative purposes without reducing effectiveness. The range restriction of the quality scores, concentrated between 6 and 9 (table 1), tends to attenuate all twelve coefficients and does not by itself explain why those for diversity are smaller. The design does not allow a decision between these explanations.

Lexical density is moderately and significantly linked to clarity and precision; with coherence the coefficient (.39; p = .060) falls at the boundary between the weak and moderate categories, and with argumentation it is weak. It is reasonable to think that texts with more content words concentrate information in fewer function words, a feature typical of written discourse, as set out in the Introduction. Because precision includes a lexical component, its association with density may reflect an overlap with the criterion, and the design does not allow a decision between these readings.

The coefficient between density and clarity requires qualification. According to the quartiles in table 1 (Q1, Mdn, and Q3 all equal to 8), half or more of the essays obtained the same clarity score, and with a null interquartile range the box plots mark as an outlier every score different from the median, so the cases in figure 2 are not necessarily extreme. The correlation rests on the minority of essays that depart from that score, whose exact number cannot be determined, and Pearson's coefficient is the most exposed to that sensitivity; a check with Spearman's coefficient would be appropriate with the individual data.

The moderate association of density with clarity and precision is consistent with the weak to moderate associations reported between lexical dimensions and scores on expository essays by Chinese learners of English (Yang et al., 2023). Differences in magnitude between studies may stem from the text genre, the participants' level, the sample size, or the way density is operationalized. With 24 participants, this study cannot settle which of these factors weighs most.

As for lexical sophistication, its association is moderate and significant with three of the four components, and the nonsignificant coefficient with clarity (.31) is not distinguishable from the others. The pattern is consistent with the finding that some second language and first language semantic context indices explained unique variance in essay assessment (Monteiro et al., 2023). The consistency is partial, because that study used semantic indices and not AWL coverage, and this result is not equivalent to a replication of that one.

AWL coverage is a frequency measure and does not inform about how appropriately each word is used. The observed median (.14) exceeds the proportion of approximately .10 that the list covers in its academic corpus (Coxhead, 2000), and token coverage may increase when a few topic-related words are repeated. Therefore, greater coverage does not equate to appropriate use of vocabulary, and the values are read as relative positions within the group; a measure that incorporated contextual appropriateness would make it possible to determine whether the association is due to the amount of academic vocabulary or to its appropriate use.

The moderate associations neither rule out nor confirm the conjecture, set out in the Introduction, that vocabulary operates together with discourse features and not as an isolated correlate of quality, because cohesion and rhetorical structure were not measured independently of the rubric. The shared variation in the significant correlations, from 18% to 23%, should be read as an optimistic upper bound: with 24 participants, coefficients that exceed the significance threshold tend to overestimate the true magnitude, and the scores depend on the judgment of three raters whose agreement the article does not report.

Essay length is another factor to consider. Total word count, an indicator of writing fluency, is among the best predictors of scores in previous studies (Appel & Arias, 2023; Yang et al., 2023). Because the task

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

targeted a length of 200 to 250 words, fluency may have varied less among the texts, although the article does not report their actual length nor was it included as a control variable, which makes it impossible to know whether part of the observed associations is due to length differences among the texts.

The longitudinal evidence on the order in which each dimension develops describes growth rates (Rokoszewska, 2024), but it does not allow inferring which dimension differentiates B1-level writers or whether that order extends to university students. Thus, the smaller magnitude of the diversity coefficients cannot be attributed, as an additional explanation to those set out, to an early stabilization of that dimension, and testing that possibility would require longitudinal data with participants at different proficiency levels.

With regard to Yang and He (2025), whose recommendations include greater exposure to low-frequency vocabulary through enriched texts, contextualized instruction, and writing tasks with feedback aimed at lexical variation, the results of this study do not allow those proposals to be assessed, since no intervention was evaluated. With weak to moderate associations for density and sophistication, negligible or weak ones for diversity, and overlapping intervals, the results do not allow a decision about which lexical features should be prioritized in teaching, a decision that would require intervention studies.

In practical terms, the results do not warrant teaching recommendations: the associations are correlational, come from 24 participants, and none survived the correction for multiple tests. The analytic rubric, however, makes it possible to identify the weaknesses of each text by component and to offer specific feedback (Weigle, 2002), and its diagnostic use does not depend on these results. Only if later studies confirmed the associations of density and sophistication with clarity and precision would it make sense to trial such feedback on academic vocabulary.

The study presents three features worth highlighting. The first is the component-by-component reading of quality, which makes it possible to describe the direction and magnitude of each association, although the independence of the components was not verified. The second is the inclusion of confidence intervals and a correction for multiple tests, which show that, with 24 participants, the associations are estimated with little precision and none survives that correction. The third is the context, English teacher education at an Ecuadorian university, where the differences between learning contexts (Zhang & Kang, 2022) justify conducting research with local samples.

5. Conclusions

The research concludes that, in this group of 24 B1-level students, the association of lexical richness with academic writing quality in English is limited and of variable magnitude. The coefficients of greatest observed magnitude correspond to density and sophistication (.30 to .48), all positive and of weak to moderate magnitude, and those for diversity range from .07 to .27, with overlapping intervals that do not allow them to be distinguished from the former. The general objective is met, since the direction, magnitude, and significance of each association were determined, its precision was estimated, and its robustness was assessed, although none survives the Bonferroni correction.

Regarding lexical diversity, the type-token ratio, calculated without correction for text length, is negligibly or weakly and nonsignificantly associated with the four quality components; the low power does not allow moderate associations to be ruled out. Lexical density shows coefficients of .30 to .48, two of them significant without correction for multiple tests (clarity and precision), and that for coherence (.39; p = .060) did not reach significance, although its magnitude is similar to that of the significant ones and the intervals overlap.

As for sophistication, academic vocabulary coverage reaches significance, without correction for multiple tests, with three components, coherence, argumentation, and precision, and shows a smaller and nonsignificant coefficient with clarity. Because the

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

confidence intervals overlap, it cannot be asserted that sophistication is associated differently with some components than with others: the pattern is described without interpreting it, and the design does not allow causal relationships to be established or the influence of other variables to be ruled out.

From a pedagogical standpoint, the data are not sufficient to either support or refute the conjecture about the joint role of vocabulary and discourse features, given that cohesion and rhetorical structure lacked their own measurement. Nor do they allow instructional interventions to be assessed, since the design was correlational and cross-sectional. If later studies that measure those features and evaluate interventions supported the conjecture, the teaching of academic vocabulary could be articulated with work on argument construction and not be limited to repertoires of isolated words.

Several limitations bound the scope of the conclusions. The small convenience sample produces unstable coefficients, prone to overestimation when they exceed the significance threshold, and not generalizable to other populations. In addition, agreement among the three raters is not reported, and Pearson's coefficient was applied to ordinal scores concentrated between 6 and 9, without checking bivariate normality. Likewise, because the twelve tests are not independent, the five significant coefficients do not constitute five separate findings.

The lexical measurement also has limits: the type-token ratio was not corrected for text length, nor was length controlled in the analysis. The three indices are not independent by construction, since the AWL is composed mostly of content words and repeated function words reduce the TTR, and their correlation is not reported. Likewise, precision includes a lexical component that may overlap with the indices, and vocabulary size and discourse features were not measured separately from the rubric.

The article does not report the data collection period, the time allowed, the actual length of the texts, or the version of VocabProfile, and the values of the lexical indices in table 1 were measured on box plots and rounded to two decimal places. In addition, two of the authors belong to the participants' program, a position whose influence could not be neutralized because the article does not report their relationship with the raters. Informed consent was verbal, which limits its later verification.

It is recommended that future research use larger, probability samples, independent raters external to the program with reported agreement, and individual data with their dispersion. It is also advisable to control text length, use diversity indices that are less sensitive to it, measure the appropriate use of academic vocabulary, and apply longitudinal designs that test the explanations offered in the Discussion.

6. References

Abu-Rabiah, E. (2023). Evaluating L2 vocabulary development features using lexical density and lexical diversity measures. *LLT Journal: A Journal on Language and Language Teaching, 26*(1), 168–182. https://doi.org/10.24071/llt.v26i1.5841

Ait Hammou, B., Larouz, M., Fagroud, M., & Akki, F. (2023). Aspects of EFL university learners' lexical and phraseological proficiency as predictors of writing quality. *Canadian Journal of Applied Linguistics, 26*(1), 91–114. https://doi.org/10.37213/cjal.2023.32726

Appel, R., & Arias, A. (2023). Automated measures of lexical sophistication: Predicting proficiency in an integrated academic writing task. *Journal of Language and Education, 9*(1), 12–25. https://doi.org/10.17323/jle.2023.11045

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology, 11*(3), 38–63. https://doi.org/10.64152/10125/44117

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8.ª ed.). Routledge.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. https://doi.org/10.2307/3587951

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5.ª ed.). SAGE.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2013). *Halliday's introduction to functional grammar* (4.ª ed.). Routledge.

Hao, Y., Jin, Z., Yang, Q., Wang, X., & Liu, H. (2023). To predict L2 writing quality using lexical richness indices: An investigation of learners of Chinese as a foreign language. *System, 118*, Article 103123.

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

https://doi.org/10.1016/j.system.2023.103123

Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill Interamericana.

Hu, R., Wu, J., & Lu, X. (2022). Word-combination-based measures of phraseological diversity, sophistication, and complexity and their relationship to second language Chinese proficiency and writing quality. *Language Learning, 72*(4), 1128–1169. https://doi.org/10.1111/lang.12511

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307–322. https://doi.org/10.1093/applin/16.3.307

Lu, X., & Hu, R. (2022). Sense-aware lexical sophistication indices and their relationship to second language writing quality. *Behavior Research Methods, 54*(3), 1444–1460. https://doi.org/10.3758/s13428-021-01675-6

Maamuujav, U., Olson, C. B., & Chung, H. (2021). Syntactic and lexical features of adolescent L2 students' academic writing. *Journal of Second Language Writing, 53*, Article 100822. https://doi.org/10.1016/j.jslw.2021.100822

Min, C., & Sukying, A. (2024). Academic vocabulary for L2 argumentative writing: Perspectives from Chinese university EFL learners. *Studies in English Language and Education, 11*(3), 1355–1373. https://doi.org/10.24815/siele.v11i3.36562

Monteiro, K., Crossley, S., Botarleanu, R.-M., & Dascălu, M. (2023). L2 and L1 semantic context indices as automated measures of lexical sophistication. *Language Testing, 40*(3), 576–606. https://doi.org/10.1177/02655322221147924

Raufi, B. S., Mulyono, H., Ilyas, H. P., & Zulaiha, S. (2024). Exploring Indonesian EFL students' lexical diversity and its correlation with academic vocabulary use in an online academic writing environment. *Ampersand, 13*, Article 100196. https://doi.org/10.1016/j.amper.2024.100196

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Rokoszewska, K. (2024). The correlations between the average semester growth rates of syntactic and lexical complexity in L2 writing development at secondary school. *Neofilolog*, (63/1), 156–177. https://doi.org/10.14746/n.2024.63.1.10

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia, 126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Susoy, Z., & Durmuşoğlu, G. (2024). An evaluation of syntactic complexity, lexical diversity and text length as predictors of foreign language writing scores. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi, 26*(4), 542–553. https://doi.org/10.17556/erziefd.1483282

Tömen, M., & Durmuşoğlu, G. (2023). Vocabulary size, lexical diversity, lexical density, and EFL writing scores: A cross-sectional study. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 13*(4), 2586–2613. https://doi.org/10.30783/nevsosbilen.1376253

Vitta, J. P., Nicklin, C., & Albright, S. W. (2023). Academic word difficulty and multidimensional lexical sophistication: An English-for-academic-purposes-focused conceptual replication of Hashimoto and Egbert (2019). *The Modern Language Journal, 107*(1), 373–397. https://doi.org/10.1111/modl.12835

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics, 168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Yang, Y., & He, X. (2025). Lexical richness in Chinese university students' EFL writing: A corpus-based comparison. *Humanities and Social Sciences Communications, 12*(1), Article 1199. https://doi.org/10.1057/s41599-025-05560-x

Yang, Y., Yap, N. T., & Mohamad Ali, A. (2023). Predicting EFL expository writing quality with measures of lexical richness. *Assessing Writing, 57*, Article 100762. https://doi.org/10.1016/j.asw.2023.100762

Yang, Y., & Zheng, Z. (2024). A refined and concise model of indices for quantitatively measuring lexical richness of Chinese university students' EFL writing. *Contemporary Educational Technology, 16*(3), Article ep513. https://doi.org/10.30935/cedtech/14707

Zhang, C., & Kang, S. (2022). A comparative study on lexical and syntactic features of ESL versus EFL learners' writing. *Frontiers in Psychology, 13*, Article 1002090. https://doi.org/10.3389/fpsyg.2022.1002090

Zheng, W. (2025). Lexical richness viewed through lexical diversity, density, and sophistication. *Digital Scholarship in the Humanities, 40*(2), 692–708. https://doi.org/10.1093/llc/fqaf023

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

Authors' biographical sketches

Ángeles María Zambrano Amores | angeles.zambrano2230@utc.edu.ec

Eighth-semester undergraduate student in the Pedagogía de los Idiomas Nacionales y Extranjeros (Pedagogy

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

of National and Foreign Languages) program at the Universidad Técnica de Cotopaxi (UTC), in Latacunga, Ecuador. Her academic research focuses on teaching and learning English as a second language. She deepens her knowledge and skills in language education, with a particular interest in effective pedagogical approaches to second language acquisition.

Ricardo Esteban Llongo Cruz | ricardo.llongo9352@utc.edu.ec

Eighth-semester undergraduate student in the Pedagogía de los Idiomas Nacionales y Extranjeros (Pedagogy of National and Foreign Languages) program at the Universidad Técnica de Cotopaxi (UTC), in Latacunga, Ecuador. His academic interests center on English as a foreign language (EFL), academic writing, lexical development, and second language acquisition. He is a coauthor of the research presented in this article on lexical characteristics and their relationship to academic writing quality in EFL contexts.

Rodrigo Vicente Tovar Viera | rodrigo.tovar@utc.edu.ec

Associate professor at the Universidad Técnica de Cotopaxi (UTC), in Latacunga, Ecuador, and holder of a PhD in Applied English Linguistics from the University of Szeged, Hungary. He completed master's degrees at the Pontificia Universidad Católica del Ecuador (PUCE) and the Universidad Indoamérica, and earned a diploma in English as a Second Language (ESL) from the Escuela de Idiomas de la ESPE and an international ESL certification from the HANSA Language Centre in Toronto. He is the author of books such as *Importance of Vocabulary Acquisition and Teaching Writing in the EFL Context* and has presented papers in Hungary, Croatia, the Czech Republic, the United States, and Ecuador.

Acknowledgments. The authors are grateful to the Universidad Técnica de Cotopaxi (UTC), the institution where the research was conducted, for the facilities provided for its completion, and to the students of the Pedagogía de los Idiomas Nacionales y Extranjeros (Pedagogy of National and Foreign Languages) program who participated in the study.

Author statements

Author contributions (CRediT taxonomy). Ángeles María Zambrano Amores: Conceptualization, investigation, data curation, formal analysis, visualization, writing (original draft). Ricardo Esteban Llongo Cruz: Conceptualization, investigation, data curation, formal analysis, visualization, writing (original draft). Rodrigo Vicente Tovar Viera: Conceptualization, methodology, validation, supervision, project administration, writing (review and editing).

Funding. This research received no external funding from public agencies, commercial entities, or nonprofit organizations.

Conflict of interest. The authors declare that they have no financial conflicts of interest that could have influenced the results or conclusions of this research. For transparency, it is reported that the authors belong to the university where the study was conducted and that two of them are eighth-semester students in the same program as the participants, one semester ahead of them. The article does not report the academic relationship between those authors and the raters, so an influence arising from that position cannot be ruled out, and it is acknowledged as a limitation in the Methodology and in the Conclusions.

Data availability statement. The article reports the medians, quartiles, and extreme values of the lexical indices,

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

DOI: https://doi.org/10.29394/Scientific.issn.2542-2987.2026.11.41.11.181-196
OAI-PMH: http://www.indteca.com/ojs/index.php/Revista_Scientific/oai

measured on box plots, and of the quality components, read from the scale of the plots (table 1), as well as the correlation coefficients and their derived statistics (tables 2 and 3). It does not report the participants' individual scores or refer to a public data repository. The analytic rubric is described in the Methodology with respect to its components and scale, without reproducing the level descriptors.

Artificial intelligence use statement. The authors declare that they used the artificial intelligence tool Claude, developed by Anthropic, as support for spelling and grammar review, for adapting the manuscript to APA 7 standards, and for verifying the correspondence between citations and references. The scoring of the essays is not among the tasks assisted by the tool, since it was carried out by three expert raters. The design of the study and the data collection were carried out by the authors, who reviewed and validated the entire document, including the interpretation of the results, and who assume full responsibility for its content.

Ethical approval and informed consent. Twenty-four university students who wrote an argumentative essay participated in the study. The study was approved by the ethics committee of the Universidad Técnica de Cotopaxi (UTC), the authors' institution, and the participants gave verbal informed consent. The results are presented in aggregate form, without information that identifies the participants.

Revista Scientific - Empirical Research Article - Peer-Reviewed - Volume 11, Issue 41 - August-October 2026 - pp. 181-196
pp. BA2016000002 - e-ISSN: 2542-2987 - ISNI: 0000 0004 6045 0361

The content of this manuscript is distributed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License