



Prosiding

Seminar Nasional Inovasi pendidikan dan Pembelajaran
Fakultas Pendidikan Bahasa dan Seni
IKIP PGRI Bojonegoro

Tema "Inovasi pendidikan dan Pembelajaran di era digital untuk Pengalaman Belajar
Imersif"



Lexical Richness Analysis in Students' Academic Writing: Implications for Vocabulary Mastery

Zalfa Shabrina¹(✉), Ayu Fitriyaningsih², Oktha Ika Rahmawati³

¹Pendidikan Bahasa Inggris, IKIP PGRI Bojonegoro, Indonesia

¹zalfashabrina077@gmail.com, ²ayu_fitriyaningsih@ikippgribojonegoro.ac.id,

³oktha_ika@ikippgribojonegoro.ac.id

abstrak—Studi deskriptif kualitatif ini bertujuan untuk mengeksplorasi kekayaan leksikal dalam tulisan akademik yang dihasilkan oleh mahasiswa semester delapan dan implikasinya terhadap penguasaan kosakata. Studi ini secara khusus menganalisis tiga dimensi kekayaan leksikal: kepadatan leksikal, keragaman leksikal, dan kecanggihan leksikal. Data dikumpulkan melalui tugas menulis terkontrol dalam bentuk esai argumentatif. Dua belas mahasiswa semester 8 berpartisipasi dalam penelitian ini. Teks-teks tersebut kemudian dianalisis secara kualitatif untuk mengidentifikasi pola penggunaan leksikal, variasi kosakata, dan tingkat kecanggihan yang ditunjukkan dalam tulisan mereka. Hasil penelitian menunjukkan bahwa kepadatan leksikal secara dominan moderat, yang menunjukkan proporsi kata isi yang seimbang dalam tulisan mahasiswa. Keragaman leksikal juga ditemukan moderat, yang menunjukkan bahwa mahasiswa menggunakan variasi kosakata yang cukup reasonable dengan beberapa pengulangan. Namun, kecanggihan leksikal tergolong rendah (rata-rata gabungan AWL dan off-list: 10,61%), yang menunjukkan bahwa mahasiswa sangat bergantung pada kosakata frekuensi tinggi dan menggunakan terbatas kata akademik atau kata maju. Temuan ini menunjukkan bahwa meskipun mahasiswa tingkat akhir menunjukkan variasi kosakata yang memadai, mereka masih perlu mengembangkan kosakata akademik mereka untuk menghasilkan tulisan akademik yang lebih sophisticated. Penelitian ini mengimplikasikan bahwa pembelajaran kosakata harus menekankan penggunaan kata akademik dan kata yang kurang sering digunakan untuk meningkatkan kecanggihan leksikal dalam tulisan mahasiswa.

Kata kunci—kekayaan leksikal, kepadatan leksikal, keragaman leksikal, kecanggihan leksikal, tulisan akademik, penguasaan kosakata

Abstract—This qualitative descriptive study aims to explore the lexical richness in academic writing produced by eighth-semester university students and its implications for vocabulary mastery. The study specifically analyzes three dimensions of lexical richness: lexical density, lexical diversity, and lexical sophistication. Data were collected through a controlled writing task in the form of an argumentative essay. Twelve semester 8 students participated in this study. The students' texts were then qualitatively analyzed to identify patterns of lexical usage, variety of vocabulary, and the level of sophistication demonstrated in their writing. The results revealed that lexical density was predominantly moderate, indicating a balanced proportion of content words in students' writing. Lexical diversity was also found to be moderate, suggesting that students used a reasonably varied vocabulary with some repetition. However, lexical sophistication was low (mean combined AWL and off-list: 10.61%), indicating that students relied heavily on high-

frequency vocabulary and used limited academic or advanced words. These findings suggest that while final-year students demonstrate adequate vocabulary variety, they still need to expand their academic vocabulary to produce more sophisticated academic writing. The study implies that vocabulary instruction should emphasize the use of academic and less frequent words to enhance lexical sophistication in students' writing.

Keywords—lexical richness, lexical density, lexical diversity, lexical sophistication, academic writing, vocabulary mastery

INTRODUCTION

Academic writing is one of the core skills that university students must master. According to Irvin in Maharani et al. (2021), academic writing is not only a means of assessment but also a form of demonstrating knowledge, critical thinking, analysis, interpretation, and presentation skills. Through writing, students are able to deliver ideas and express it in a written way (Nurdianingsih & Purnama, 2017; Widyaningrum et al., 2022). However Juniarti (2019) further states that producing a good academic text requires creativity, initiative, and the ability to organize ideas productively so that readers can understand the writer's intended message. Therefore, academic writing ability is often regarded as an important benchmark of students' success in higher education.

One of the most significant factors influencing the quality of academic writing is vocabulary mastery. Vocabulary serves as the foundation for constructing meaningful sentences, paragraphs, and texts (Estika et al., 2025). It forms the foundation for all other language skills, including reading, writing, speaking, and listening (Rochmadhoni et al., 2022; Aprilianti et al., 2025). The richer and more varied a writer's vocabulary, the greater their ability to express ideas with appropriate nuance, precision, and depth (Choemue & Bram, 2021). Thus, vocabulary mastery is a crucial indicator of writing quality.

To measure vocabulary mastery objectively, lexical richness analysis has been widely used by researchers. This study adopts three main dimensions of lexical richness: lexical density, lexical diversity, and lexical sophistication. Lexical density refers to the proportion of content words (nouns, verbs, adjectives, and adverbs) to the total number of words in a text (Kyle, 2020; Read in Zhang et al., 2021). Lexical diversity measures the variety of different words used in the writing, where a higher number of unique words indicates greater diversity (Johansson in Choemue & Bram, 2021; McCarthy & Jarvis, 2010). Lexical sophistication, on the other hand, concerns the use of low-frequency, advanced, and academic vocabulary (Signes & Arroitia in Kusumanagrum & Ardi, 2019; Juanggo, 2018). A higher level of lexical richness generally indicates a broader and deeper vocabulary repertoire, which in turn reflects stronger vocabulary mastery. Therefore, analyzing these three dimensions is essential for understanding learners' overall vocabulary development and their ability to use vocabulary effectively in written discourse.

However, although many studies have examined these dimensions, research that comprehensively integrates lexical density, diversity, and sophistication while specifically exploring their implications for students' vocabulary mastery is still limited. Most previous studies tend to investigate these aspects separately (e.g., Zeng, 2021; Juanggo, 2018; Meabangsai et al., 2023) or only present descriptive findings without deeply linking them to the broader construct of vocabulary

development in academic writing. Addressing this gap, the present study aims to analyze the lexical richness of students' academic writing by examining the three dimensions simultaneously to provide a deeper insight into their vocabulary mastery. The results are expected to offer implications for understanding how these lexical dimensions support more effective and qualified academic writing.

RESEARCH METHODOLOGY

This study employed a descriptive qualitative design to analyze the lexical richness of students' argumentative essays and its implications for vocabulary mastery in academic writing. The research was conducted at IKIP PGRI Bojonegoro, specifically in the English Education Department of the Faculty of Languages and Arts Education. The participants were twelve eighth-semester students from two classes in the English Education Department, with an estimated total population of around 60 students. Although these students were not currently enrolled in a writing course, they had previously completed four writing courses: Paragraph Writing; Narrative & Descriptive Writing; Expository & Argumentative Writing; and Research Paper Writing. The participants wrote their essays by hand on paper during supervised data-collection sessions, and consent for the use of their data was collected through a Google Form.

A purposive sampling strategy with authenticity screening was used. Students were invited to write an argumentative essay in response to the prompt: *"Do you agree or disagree that university students should have part-time jobs while studying? Give reasons and examples to support your opinion."* The essays were written directly on paper during supervised sessions within two weeks in the second semester of the 2025/2026 academic year. The inclusion criteria were: (1) eighth-semester student, (2) essay written during the supervised session, (3) essay in English, and (4) participant consent. Essays were excluded if they were typed or previously submitted, suspected of involving AI or third-party assistance, or incomplete. Twelve essays met all inclusion criteria and were selected as the final corpus. Given the descriptive qualitative design, the emphasis on data authenticity (handwritten, on-the-spot writing), and time constraints, 12 essays were considered sufficient for in-depth qualitative description of lexical richness patterns.

The data collection process began by collecting the students' handwritten essays. After that, the essays were transcribed verbatim into digital text for analysis. The texts were then checked and prepared before being analyzed to ensure that the words, sentences, and punctuation were readable and suitable for processing. Lexical richness was examined through three dimensions: lexical density, lexical diversity, and lexical sophistication, and were analyzed using VocabProfiler software. Lexical density was calculated as the ratio of content words to total words. Lexical diversity was measured using the Type-Token Ratio (TTR). Lexical sophistication was analyzed using the Lexical Frequency Profile (LFP) to identify the proportion of advanced or less frequent vocabulary. After the scores were obtained, the results were interpreted descriptively to describe the students' lexical richness.

RESEARCH FINDINGS & DISCUSSION

The analysis of lexical richness in the 12 argumentative essays revealed patterns in three dimensions: lexical density, lexical diversity, and lexical sophistication. To interpret the results, the following categories were used: lexical density was classified as low (< 40%), moderate (40–50%), and high (> 50%); lexical diversity (TTR) as low (< 0.40), moderate (0.40–0.60), and high (> 0.60); and lexical sophistication based on the combined proportion of off-list and Academic Word List (AWL) words as low (< 15%), moderate (15–25%), and high (> 25%).

The following sections discuss each dimension of lexical richness in more detail, followed by the implications for vocabulary mastery.

Table 1. Lexical richness results

Essay	Density	Diversity (TTR)	AWL	Off-list	Sophistication
S01	45%	0.42	9.68 %	3.23%	12.91%
S02	45%	0.47	2.89 %	1.73%	4.62%
S03	61%	0.46	3.53 %	7.65%	11.18%
S04	43%	0.51	3.54 %	5.05%	8.59%
S05	55%	0.61	8.33 %	2.60%	10.93%
S06	46%	0.48	3.86 %	2.81%	6.67%
S07	51%	0.56	5.33 %	5.33%	10.66%
S08	58%	0.43	9.20 %	3.48%	12.68%
S09	43%	0.44	6.25 %	5.77%	12.02%
S10	59%	0.49	7.48 %	2.04%	9.52%
S11	49%	0.42	2.97 %	3.96%	6.93%
S12	54%	0.48	6.02 %	8.03%	14.05%
mean	50%	0.48	6.30 %	4.31%	10.61%

Lexical Density

The lexical density of the students' writing was found to be moderate with an average of 50.0%. This indicates that only about half of the words used were content words, while the rest were function words. This pattern suggests that the students

frequently relied on basic connective words to link ideas rather than employing richer and more varied vocabulary. A clear example can be seen in the following excerpt:

"I can't focus on my assignment in the college because my body was so exhausted and if I do my task in my working place, I can't do it properly because I was thinking my assignment in the college."

The repeated use of words such as *because*, *and*, *if*, *in*, and *my* contributes to the lower lexical density. Such patterns were commonly found across the participants' texts, indicating that many students still struggle to produce lexically dense academic writing.

Lexical Diversity

Lexical diversity in the students' writings was generally at a moderate level, with an average Type-Token Ratio (TTR) of 0.48. This score indicates that, on average, nearly half of the words used in the essays were unique words, while the other half consisted of repetitions of previously used words.

Although some students demonstrated relatively better lexical diversity (with TTR reaching up to 0.61), the majority of the participants showed moderate to low diversity. This suggests that students tended to repeat the same set of words throughout their essays rather than using a wide range of vocabulary to express their ideas. Common words such as *money*, *job*, *student*, *study*, *work*, *because*, and *need* were repeatedly used by most participants.

For example, one student repeatedly emphasized the word "money" in their writing:

"money money money i like money... everything in life is about money"

This repetition reflects limited lexical variation. The moderate lexical diversity found in this study implies that even though the students are in their final semester, they still rely heavily on a narrow range of familiar vocabulary when discussing familiar topics. This limitation reduces the richness and sophistication of their academic writing.

Lexical Sophistication

Lexical sophistication represents the use of low-frequency and advanced vocabulary, particularly academic words. In this study, the students demonstrated low lexical sophistication, as indicated by the average combined percentage of Academic Word List (AWL) and Off-List words of only 10.61%. This means that more than 85% of the vocabulary used by the students still belonged to the most basic 1,000 high-frequency words (K1).

This low level of sophistication shows that the participants rarely employed more advanced or academic vocabulary in their argumentative essays. Most students relied heavily on everyday high-frequency words such as *money*, *job*, *student*, *study*, *work*, *need*, and *important*. Even when discussing topics related to part-time jobs and university life, they seldom used more precise academic terms such as financial independence, time management, opportunity, responsibility, or prioritization. This is evident in sentences such as:

"Actually we are need a money, but for student i think is not a obligation, because we need to focus for study and have a great score."

"..with money we are happy, have a good mood, one more time money is a save everything."

These sentences use very basic vocabulary and contain grammatical errors, with no use of academic or less frequent words. This illustrates the limited range of vocabulary and the lack of lexical sophistication. Although a few students showed slightly better performance (with sophistication index above 13%), the overall finding reveals that final-year students still face significant challenges in using varied and advanced vocabulary in their academic writing.

Implications for Vocabulary Mastery

The findings of this study reveal that although the participants are eighth-semester students who are nearing graduation, their lexical richness in academic writing remains relatively limited. With an average lexical density of 50%, moderate lexical diversity (TTR = 0.48), and notably low lexical sophistication (10.61%), this study suggests that the students' vocabulary mastery is still underdeveloped, particularly in producing sophisticated academic texts.

These results imply that many students have not yet achieved a strong productive vocabulary mastery despite years of English learning. They tend to rely on basic, high-frequency words, and struggle to use more advanced academic vocabulary and varied expressions. This limitation may hinder their ability to convey ideas effectively in academic and professional contexts after graduation.

The findings highlight two important implications for vocabulary instruction in higher education. First, explicit vocabulary teaching is still needed even at the advanced level. Teachers should focus more on academic word lists (AWL) and encourage students to use low-frequency words appropriately. Second, writing practice alone is not sufficient; students need systematic vocabulary development activities, such as synonym practice, word family exercises, and exposure to academic texts. Therefore, integrating vocabulary enhancement strategies into content courses is highly recommended.

Overall, this study underscores the need for more intensive and deliberate vocabulary instruction throughout the undergraduate program to help students develop richer and more sophisticated lexical resources before entering the workforce.

CONCLUSION

This study concludes that the students' academic writing showed moderate lexical density, moderate lexical diversity, and low lexical sophistication. These results indicate that the students were able to use a reasonably varied range of vocabulary, but they still relied mostly on high-frequency words and used only a limited number of academic or advanced terms. In other words, their vocabulary mastery was not yet strong enough to support highly sophisticated academic writing.

Therefore, the findings suggest that vocabulary instruction in higher education should pay more attention to the use of academic and less frequent words. Students

also need more systematic practice to expand their productive vocabulary so they can express ideas more precisely and produce better academic texts.

REFERENCE

- Aprilianti, E., Tirtanawati, M. R., & Prastiwi, C. H. W. (2025). Using Flashcard to Enhance the Vocabulary of the Fourth Graders of Elementary School. In *Seminar Nasional Inovasi pendidikan dan Pembelajaran*, 1(01), 301-307.
- Choemue, S., & Bram, B. (2021). Lexical Richness in Scientific Journal Articles A Comparison between ESL and EFL Writers. *Indonesian Journal of EFL and Linguistics*, 6(1), 147. <https://doi.org/10.21462/ijefl.v6i1.349>.
- Estika, A. L., Tirtanawati, M. R., & Prastiwi, C. H. W. (2025). Improving Students Vocabulary Mastery Through Tpr (Total Physical Response) Assisted by Song Media. In *Seminar Nasional Inovasi pendidikan dan Pembelajaran* (Vol. 1, No. 1, pp. 442-447).
- Juanggo, W. (2018). Investigating lexical diversity and lexical sophistication of productive vocabulary in the written discourse of Indonesian EFL learners. *Indonesian Journal of Applied Linguistics*, 8(1), 38-48. <https://doi.org/10.17509/ijal.v8i1.11462>.
- Juniarti, Y. (2019). Pentingnya keterampilan Menulis Akademik di
- Kusumaningrum, M. V., & Ardi, P. (2019). *A Corpus Analysis of Lexical Sophistication in Litjournal: A Journal on Language and Language Teaching*.
- Kyle, K. (2020). Measuring Lexical Richness. In *The Routledge Handbook of Vocabulary Studies*.
- Maharani, D. P., Rukmini, D., Faridi, A., & Astuti, P. (2021). Students' Academic Writing Competence Realized in Writing Research Proposals: Viewed from their Content, Organization and Cohesion, Range, Register, and the Accuracy of Language. *Issue I Ser. II*, 11, 36-43. <https://doi.org/10.9790/7388-1101023643>.
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392. <https://doi.org/10.3758/BRM.42.2.381>.
- Meebangsai, D., Pongtin, P., Kitipoontanakorn, P., & Laosrirattanachai, P. (2023). *Investigating Proficiency of Academic English in Student Writing: A Comparative Case Study on Vocabulary Utilization in Student Research Article Writing vis-à-vis National and International Research* (Vol. 67). <https://digital.car.chula.ac.th/pasaa>.
- Nurdianingsih, F., & Purnama, Y. I. (2017). Thematic progression pattern: A technique to improve students' writing skill viewed from writing apprehension. *Journal of Linguistic and English Teaching*, 2(2), 237-247.
- Perguruan Tinggi. In *Prosiding Sembadra Universitas Sriwijaya* (Vol. 2, Issue 1).
- Rochmadhoni, N., Prastiwi, C. H. W., & Rahmawati, O. I. (2022). The effectiveness of crossword puzzle games in teaching vocabulary to the seventh graders at SMPN 2 Bojonegoro. *ELTT*, 8(1), 99-103.
- Widyaningrum, H. K., Pratiwi, C. P., Menggala, A. D., Hasanudin, C., & Fitriyaningsih, A. (2022). Android application appy pie to support students writing stories skill through flipped classroom learning models. *International*

Journal on Advanced Science, Engineering and Information Technology, 12(2), 530-538.

Zeng, M. (2021). The Influence of Lexical Features on Human Rater's Judgement of Essay Performance on Local English Placement Test.

Zhang, H., Chen, M., & Li, X. (2021). Developmental Features of Lexical Richness in English Writings by Chinese Beginner Learners. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.665988>.