

**LEXICAL RICHNESS ANALYSIS IN STUDENTS' ACADEMIC
WRITING: IMPLICATIONS FOR VOCABULARY MASTERY**

SKRIPSI



**In partial fulfilment of the requirements
For the degree of Sarjana in English Education**

By:

ZALFA SHABRINA

22120064

**ENGLISH EDUCATION DEPARTMENT
FACULTY OF LANGUAGES AND ARTS EDUCATION
IKIP PGRI BOJONEGORO**

2026

**LEXICAL RICHNESS ANALYSIS IN STUDENTS' ACADEMIC
WRITING: IMPLICATIONS FOR VOCABULARY MASTERY**

SKRIPSI

Presented to

IKIP PGRI Bojonegoro

in partial fulfilment of the requirements
for the degree of Sarjana in English Education

By:

ZALFA SHABRINA

NIM 22120064

**ENGLISH EDUCATION DEPARTMENT
FACULTY OF LANGUAGES AND ARTS EDUCATION
IKIP PGRI BOJONEGORO**

2026

APPROVAL SHEETS

The skripsi entitled "**Lexical Richness Analysis in Students' Academic Writing:
Implication for Vocabulary Mastery**" is compiled by:

Name : Zalfa Shabrina
NIM : 22120064
Department : English Education
Has been approved by the advisors.

Bojonegoro, 7 July 2026

Advisor I,



Ayu Fitrianiingsih, S.Pd., M.Pd.
NIDN: 0720049101

Advisor II,



Oktha Ika Rahmawati, S.Pd., M.Pd.
NIDN: 0701108602

LEGITIMATION

The skripsi entitled "LEXICAL RICHNESS ANALYSIS IN STUDENTS' ACADEMIC WRITING: IMPLICATIONS FOR VOCABULARY MASTERY" is compiled by:

Name : Zalfa Shabrina
Nim : 22120064
Department : English Education

This is to certify that this skripsi has been approved by the Board of Examiners as the requirement for the degree of Sarjana in English Education Department, Faculty of Languages and Arts Education, IKIP PGRI Bojonegoro on 23 July 2026

Bojonegoro, 23 July 2026

Chairman,



Dr. Cahyo Hasanudin, M.Pd.
NIDN 0706058801

Secretary,



Ayu Fitrianiingsih, M.Pd.
NIDN 0720049101

Examiner I,



Dr. Moh. Fuadul Matin, S.S., M.Pd.
NIDN 0727028703

Examiner II,



Chyntia Heru Woro Prastiwi, M.Pd.
NIDN 0728017903

Rector,

Dr. Junarti, M.Pd.
NIDN 0014016501

MOTTO

“Love what you like, it’ll bloom way more beautifully than you’ve ever known.”

– Zalfa Shabrina

“It’s such a great story, isn’t it? Let’s meet again, Yoo Joonghyuk.”

– Kim Dokja in chapter 188: Ep. 35 – The 73rd Demon King, VII

ACKNOWLEDGEMENT

First and foremost, all praise and deepest gratitude go to Allah SWT, the One and Only God, who has always blessed me with good health, strength, and endless ease. Thanks to His guidance, I was able to complete this thesis entitled **“Lexical Richness Analysis in Students’ Academic Writing: Implications for Vocabulary Mastery”** after such a long journey.

I would also like to thank the entire IKIP PGRI Bojonegoro family, the university administration, all the lecturers and the supportive staff for the knowledge, facilities, and constant guidance throughout my studies.

My sincere appreciation also goes to my two advisors, Mrs. Ayu Fitrianiingsih, S.Pd., M.Pd. and Mrs. Oktha Ika Rahmawati, S.Pd., M.Pd., who have patiently guided, supported, and motivated me from the beginning until the end of this skripsi. Their valuable advice, constructive feedback, and encouragement have been incredibly meaningful to me.

In addition, I want to express my heartfelt thanks to the following amazing people who have given me extraordinary help and moral support:

1. To my beloved parents, who have always supported my every wish and need with endless love and sincere prayers. Thank you for your endless sacrifices, for the constant meals prepared with care, for the encouragement, and for the strength you gave me when I felt like giving up. Your unwavering belief in me was the foundation that carried me through my studies. This achievement is not mine alone, it is the result of your love, patience, and prayers that never once stopped accompanying me. My deepest gratitude will always be yours.

2. To my dearest friend, Na'ma'ul Jazilah, who has patiently listened to all my complaints and silly jokes without ever getting tired. Your comforting presence and genuine support have been a great source of strength during the challenging moments of this thesis journey.
3. To my dear classmates: Ma'rifatul Ain, Rochma Annisa Choiriyah, and Dyah Arum Centhini Tambang Raras, for all the moments and support we shared. Thank you for the laughter, discussion, and mutual encouragement that made this journey more meaningful and enjoyable.
4. To my fellow English Education major, who have accompanied and helped me through many parts of this thesis journey.
5. To my dear childhood friend, Imroatus Sholihah, thank you for your constant support and all the beautiful memories we've created together over the years. Thank you for always sharing everything that happens out there and for helping me with so many ideas while working on this thesis.
6. To my beloved chickens, especially Pipip, Piyik, and the whole crew, thank you for being my little bundle of joy and emotional support. Your cheerful and chirpy presence has helped me tremendously, particularly during the toughest times when I felt overwhelmed. You made completing this thesis feel much more bearable.
7. Special thanks to Kim Dokja from Omniscient Reader's Viewpoint, who has been my greatest source of mental support. His story, unwavering spirit, and perseverance in facing countless challenges have deeply inspired me. He motivated me to keep going and complete this thesis with even stronger determination.

I realize that this thesis is still far from perfect. Therefore, I would greatly appreciate any constructive criticism, suggestions, and feedback from all parties to improve this work.

Bojonegoro, 7 July 2026
The Researcher

Zalfa Shabrina
NIM 22120064

STATEMENT OF AUTHORSHIP

The undersigned below:

Name : Zalfa Shabrina
NIM : 22120064
Department : English Education
Faculty : Languages and Arts Education

In order to uphold academic integrity, sincerely and without coercion from any parties, I hereby declare that the thesis entitled:

LEXICAL RICHNESS ANALYSIS IN STUDENTS' ACADEMIC WRITING: IMPLICATIONS FOR VOCABULARY MASTERY

Is the result of my own original work, and all sources of information used have been clearly cited in the reference list according to the scientific and academic code of ethics. I am aware that if any violations of scholarly ethics regarding to the authenticity of this work are found, **I personally** accept the consequences in accordance with the applicable regulations and am prepared to bear legal sanctions.

Bojonegoro, 7 July 2026
The Researcher



Zalfa Shabrina
NIM 22120064

ABSTRACT

Shabrina, Zalfa. 2026. Lexical Richness Analysis in Students' Academic Writing: Implications for Vocabulary Mastery. Skripsi, English Education Department, Faculty of Languages and Arts Education, IKIP PGRI Bojonegoro, Advisor (I) Ayu Fitrianingsih, S.Pd., M.Pd., (II) Oktha Ika Rahmawati, S.Pd., M.Pd.

This study aimed to analyze the lexical richness of students' academic writing and to explore its implications for their vocabulary mastery. Lexical richness was examined through three dimensions: lexical density, lexical diversity, and lexical sophistication. This research employed an explanatory sequential mixed-methods design, in which quantitative measurement was followed by qualitative textual analysis to provide a more detailed interpretation of the findings. The participants were 20 eighth-semester students of the English Education Department at IKIP PGRI Bojonegoro, selected through purposive sampling. Data were collected through a supervised, timed argumentative essay writing task and analyzed using VocabProfiler for lexical density and lexical sophistication, and TAALED (MATTR) for lexical diversity. The findings revealed that the students' academic writing was characterized by low lexical density ($M = 49.05\%$, $SD = 0.062$), relatively high lexical diversity (MATTR $M = 0.728$, $SD = 0.055$), and very low lexical sophistication ($M = 8.319\%$, $SD = 0.026$). Further qualitative analysis showed that the students' apparent lexical diversity was largely produced through variation among high-frequency, everyday words rather than through the use of academic or lower-frequency vocabulary. This pattern indicates that the students possess considerable breadth of productive vocabulary but have not yet developed its depth, particularly for academic purposes. These findings suggest that students' vocabulary mastery in academic writing remains limited to general communicative use, and that explicit, register-focused academic vocabulary instruction is needed to help students, especially as prospective English teachers, produce more lexically sophisticated academic writing.

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

ABSTRAK

Shabrina, Zalfa. 2026. Lexical Richness Analysis in Students' Academic Writing: Implications for Vocabulary Mastery. Skripsi, English Education Department, Faculty of Languages and Arts Education, IKIP PGRI Bojonegoro, Advisor (I) Ayu Fitrianingsih, S.Pd., M.Pd., (II) Oktha Ika Rahmawati, S.Pd., M.Pd.

Penelitian ini bertujuan untuk menganalisis kekayaan leksikal dalam tulisan akademik mahasiswa serta mengeksplorasi implikasinya terhadap penguasaan kosakata mereka. Kekayaan leksikal dikaji melalui tiga dimensi, yaitu kepadatan leksikal, keragaman leksikal, dan kecanggihan leksikal. Penelitian ini menggunakan desain mixed-methods eksplanatori sekuensial, di mana pengukuran kuantitatif dilakukan terlebih dahulu, kemudian dilanjutkan dengan analisis tekstual kualitatif untuk memberikan interpretasi yang lebih mendalam terhadap temuan. Partisipan penelitian ini adalah 20 mahasiswa semester delapan Program Studi Pendidikan Bahasa Inggris IKIP PGRI Bojonegoro yang dipilih melalui purposive sampling. Data dikumpulkan melalui tugas menulis esai argumentatif yang dilakukan secara terjadwal dan diawasi, kemudian dianalisis menggunakan VocabProfiler untuk mengukur kepadatan dan kecanggihan leksikal, serta TAALED (MATTR) untuk mengukur keragaman leksikal. Hasil penelitian menunjukkan bahwa tulisan akademik mahasiswa memiliki kepadatan leksikal yang rendah ($M = 49,05\%$, $SD = 0,062$), keragaman leksikal yang relatif tinggi (MATTR $M = 0,728$, $SD = 0,055$), dan kecanggihan leksikal yang sangat rendah ($M = 8,319\%$, $SD = 0,026$). Analisis kualitatif lebih lanjut menunjukkan bahwa keragaman leksikal yang tampak tersebut sebagian besar dihasilkan dari variasi kata-kata umum berfrekuensi tinggi, bukan dari penggunaan kosakata akademik atau kosakata berfrekuensi rendah. Pola ini mengindikasikan bahwa mahasiswa telah memiliki keluasan (breadth) kosakata produktif yang cukup baik, tetapi belum diikuti dengan kedalaman (depth) kosakata, terutama untuk keperluan akademik. Temuan ini menunjukkan bahwa penguasaan kosakata mahasiswa dalam menulis akademik masih terbatas pada penggunaan komunikatif umum, sehingga diperlukan pengajaran kosakata akademik yang eksplisit dan berorientasi pada register tertentu agar mahasiswa, khususnya sebagai calon guru bahasa Inggris, mampu menghasilkan tulisan akademik yang lebih kaya secara leksikal.

Kata kunci: kekayaan leksikal, kepadatan leksikal, keragaman leksikal, kecanggihan leksikal, penguasaan kosakata, penulisan akademik

PREFACE

All praise and gratitude are due to Allah SWT for His endless blessings and guidance, without which I would not have been able to complete this thesis entitled “Lexical Richness Analysis in Students’ Academic Writing: Implications for Vocabulary Mastery”. The process of completing this thesis was not easy. There were moments of confusion, fatigue, and self-doubt. However, through all the challenges, I learned a lot not only about academic writing but also about patience, perseverance, and the importance of continuous learning. I hope that this study can provide useful insights for English learners, fellow students, and lecturers in understanding the importance of vocabulary development in academic writing. I also hope this thesis can serve as a small contribution to the field of English language teaching.

Finally, I am aware that this thesis is still far from perfect. Therefore, I sincerely welcome any constructive criticism and suggestions for its improvement.

Bojonegoro, 7 July 2026
The Researcher

Zalfa Shabrina
NIM 22120064

TABLE OF CONTENTS

TITLE PAGE	ii
APPROVAL SHEETS.....	iii
LEGITIMATION.....	iv
MOTTO	v
ACKNOWLEDGEMENT.....	vi
STATEMENT OF AUTHORSHIP	ix
ABSTRACT	x
ABSTRAK	xi
PREFACE.....	xii
TABLE OF CONTENTS.....	xiii
LIST OF TABLES	xv
LIST OF FIGURES	xvi
LIST OF APPENDICES	xvii
CHAPTER I.....	1
INTRODUCTION.....	1
A. Background of Study	1
B. Formulation of the Problems.....	4
C. Objectives of the Study	4
D. Significances of the Study.....	5
E. Definon of Terminologies.....	6
CHAPTER II	9
REVIEW OF RELATED LITERATURE.....	9
A. Review of Previous Studies	9
B. Theoretical Review	14
C. Theoretical Framework	32
D. Research Hypothesis	34
CHAPTER III	35
RESEARCH METHODOLOGY	35
A. Research Design.....	35
B. Research Procedures	35
C. Research Settings	36

D. Population, Sample, and Sampling	37
E. Data Collection Techniques	38
F. Data Analysis Techniques.....	39
G. Data Validation Techniques.....	40
CHAPTER IV.....	42
RESEARCH FINDINGS AND DISCUSSION.....	42
A. Findings.....	42
B. Discussions.....	55
CHAPTER V	67
CONCLUSION AND SUGGESTION	67
A. Conclusion	67
B. Suggestion.....	68
REFERENCES.....	71
APPENDIX	75

LIST OF TABLES

Table 4. 1 Lexical Density of Students' Argumentative Essays	43
Table 4. 2 Statistics of Lexical Density	44
Table 4. 3 Lexical Diversity (MATTR) of Students' Argumentative Essays.....	47
Table 4. 4 Statistics of Lexical Diversity (MATTR)	47
Table 4. 5 Lexical Sophistication of Students' Argumentative Essays	50
Table 4. 6 Statistics of Lexical Sophistication	51

LIST OF FIGURES

Figure 2. 1 Theoretical Framework.....	32
--	----

LIST OF APPENDICES

Appendix 1 The Informed Consent for Data Use	75
Appendix 2 Students' Argumentative Essays	80
Appendix 3 Documentation of Students' Writing Essay	101

CHAPTER I

INTRODUCTION

A. Background of Study

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 key factors determining the quality of academic writing is vocabulary mastery. Vocabulary serves as the pondation for constructing sentences, paragraphs, and entire texts (Estika et al., 2025). As Liem & Suryo (2025) state that a learner cannot understand texts, communicate ideas verbally or in writing, or participate in conversations without an adequate vocabulary range. Therefore, vocabulary is considered the foundation for all other language skills—reading, writing, speaking, and listening (Rochmadhoni et al., 2022;Aprilianti et al., 2025). This means that the more varied a writer's vocabulary, the greater the ability to express ideas with appropriate nuance, variety, and depth. This aligns with Mufidah et al. (2020) that learners with a

broad, well-established vocabulary comprehend language more easily and can express themselves better in speech, writing, and reading; consequently, their writing tends to be more varied and less monotonous. In other words, vocabulary mastery is an important indicator for assessing the quality of students' writing.

To assess vocabulary mastery objectively, one effective approach is lexical richness analysis. In the literature, the operationalization of this construct is varied. For instance, Laufer and Nation in Zhang et al. (2021) originally defined lexical richness consists of four dimensions: lexical density, lexical diversity, lexical sophistication, and lexical error. However, contemporary research often focuses on the productive aspects of vocabulary rather than error analysis. Kyle (2020), for example, measures lexical richness using only three core aspects: density, diversity, and sophistication. Similarly, Choemue & Bram (2021) utilized this three-dimensional framework to compare the writing of ESL and EFL learners. Aligning with these studies, the present research excludes lexical error and concentrates exclusively on lexical density, diversity, and sophistication to provide a comprehensive measure of vocabulary richness. The specific definitions of these three distinct but interrelated components are explained below.

Johansson in Choemue & Bram (2021), explains that lexical density generally refers to the proportion of content words—nouns, verbs, adjectives, and adverbs—compared to the total number of words in a text. On the other hand, lexical diversity relates to the range of vocabulary used, namely the number of different types of words produced in a particular text (Juanggo,

2018). Based on these concepts, Choemue & Bram (2021) describe lexical sophistication as the extent to which writers use “unusual” or more advanced words in their writing. Overall, lexical density, diversity, and sophistication provide a comprehensive picture of a learner’s lexical richness as lexical density and diversity provide quantity of the vocabulary used, while lexical sophistication provides the quality of the vocabulary used. 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, despite the importance of these measures, research that integrates all three aspects—density, diversity, and sophistication—to specifically explore their collective implications for vocabulary mastery remains limited. Most previous studies tend to focus on these dimensions separately or only as descriptive statistics without linking them directly to the broader construct of mastery. Addressing this gap, this research is designed to analyze students’ lexical richness in academic writing comprehensively. By examining lexical density, diversity, and sophistication simultaneously, this research aims to provide deeper insight into the extent of students’ vocabulary mastery. Ultimately, the results will offer significant implications for understanding how these lexical dimensions support more effective and qualified academic writing.

B. Formulation of the Problems

This research addresses the following research questions:

1. What is the lexical density of students' academic writing measured by the proportion of content words relative to total words?
2. What is the lexical diversity of students' academic writing measured by MATTR?
3. What is the lexical sophistication of students' academic writing measured by LFP?
4. What are the implications of these lexical richness dimensions for students' vocabulary mastery in academic writing?

C. Objectives of the Study

This research aims to achieve the following objectives:

1. To analyze the lexical density of students' academic writing as measured by the proportion of content words relative to total words.
2. To analyze the lexical diversity of students' academic writing as measured by MATTR.
3. To analyze the lexical sophistication of students' academic writing as measured by sophisticated word proportion and AWL coverage.
4. To explore the implications of these lexical richness dimensions for students' vocabulary mastery in academic writing.

D. Significances of the Study

1. Theoretical Significances

This research provides several important theoretical contributions related to the relationship between lexical profiles and vocabulary mastery in academic writing:

a. Refinement of the vocabulary richness dimension

Establishing density, diversity (MATTR), and complexity as validated indicators of productive vocabulary mastery in academic writing in English as a foreign language (EFL).

b. Validation of MATTR

Providing empirical validation of MATTR as a superior measure of lexical diversity, establishing benchmarks for analyzing student writing.

c. Lexical sophistication framework

Demonstrating the use of sophisticated words (beyond 2K + AWL) as a key predictor of academic proficiency.

d. Integrated development model

Developing a theoretical model that links lexical richness with vocabulary mastery and rhetorical competence in a productive context.

2. Practical Significances

This research offers several practical contributions from the analysis of lexical richness' aspects (density, diversity, sophistication) and their implications for vocabulary mastery.

a. For lecturers

Providing multidimensional lexical indicators (density, MATTR, complexity level) for accurate diagnosis and targeted feedback on students' vocabulary weaknesses.

b. For students

Providing concrete feedback on lexical aspects that need improvement (e.g., increasing word variety, reducing repetition) allows students to practice with clear and focused goals.

c. For developing teaching materials

Providing an empirical basis for designing specific vocabulary exercises aligned with the demands of academic writing.

d. For future researchers

Providing a verified data set for intervention testing and further metric comparison in studies of English vocabulary as a foreign language.

E. Definiton of Terminologies

1. Lexical Richness

Hoover in Choemue & Bram (2021) defined lexical richness as the ratio of types of words to the total words (token) written in a text.

a. Lexical density

Lexical density is a measurement of density by comparing the number of content words to the total number of words in a text (Kyle, 2020).

Read and Biber explain in Zhang et al. (2021) study that lexical density, defined as “the proportion of lexical words in a text,” reflects the information load of the text. Johansson in Choemue & Bram

(2021) stated that lexical density mainly refers to the number of content words, such as nouns, verbs, adjectives, and adverbs, compared to the total number of words in the text.

b. Lexical diversity

Signes & Arroitia in Kusumaningrum & Ardi (2019) stated that lexical diversity deals with how many different words are used in the text. McCarthy & Jarvis (2010) define lexical diversity as the range of different words used in a text, with a wider range indicating greater diversity. Juanggo (2018) said lexical diversity or lexical variation is a measure of how extensive the vocabulary used in a text is (how many different words are used).

c. Lexical sophistication

Signes & Arroitia in Kusumaningrum & Ardi (2019) stated that lexical sophistication refers to the portion of advanced vocabulary. Read in Zhang et al. (2021) explained that Lexical sophistication measurement refers to “the proportion of relatively uncommon or sophisticated words in texts written by learners.” Choemue & Bram (2021) stated that lexical sophistication is a measure that aims to determine how many words that are considered “uncommon” or sophisticated are used by writers in their writing.

2. Academic Writing

According to USC Libraries (University of Southern California), academic writing is an expression to represents the intellectual boundaries of disciplines, characterized by a formal tone, specialized terminology,

third-person perspective, concise focus, and precise word choice, all intended to communicate complex ideas clearly within scholarly communities. Adri (2024) defined academic writing as a specialized form of writing in higher education that requires mastery of both the subject matter and the target language.

3. Vocabulary Mastery

Schmitt in Rahmah et al. (2023) defined vocabulary mastery as the ability to understand and command words effectively. Such mastery plays a crucial role in English learning, since it enables learners to access and expand their knowledge base. Safi'i (2022) defines vocabulary mastery as the number of words that a student recognizes and has a good understanding of, which can be measured, for example, through vocabulary tests that require the selection of synonyms, similar meanings, or equivalents for given words.