

People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

University of Guelma- 8 Mai 1945

Faculty of Letters and Languages

Department of English

جامعة 8 ماي 1945 قالمة

كلية الآداب واللغات

قسم اللغة الإنجليزية



A Dissertation Submitted in Partial Fulfillment of the Requirements of Master's Degree
in Language and Culture

Title:

**Evaluating Lexical Richness in Higher Education: A Comparative
Study through the Use of AI**

Submitted by:

-Ms. Setiti Rania

-Ms. Boulmakh Insaf Hadia

Supervised by:

Dr. Ilyes Gouider

Board of Examiners:

-Chair: Dr. Mahfoud Ali Zoui (MCA)

University of 8 May 1945-Guelma

-Supervisor: Dr. Ilyes Gouider (MAB)

University of 8 May 1945-Guelma

-Examiner: Mrs. Yamina Biskri (MAA)

University of 8 May 1945-Guelma

Academic Year: 2024/2025

Abstract

The present research aims at highlighting the efficiency of artificial intelligence (AI) and Natural Language Processing tools (NLP) in text analysis and the assessment of students' language proficiency. The study targets one dimension of language proficiency embodied in the lexical richness of students' writings. It represents a study focused on measuring the density, sophistication, and diversity of students' vocabulary. A descriptive-correlational research design was adopted while employing a quantitative method for data analysis. The participants involved 60 English as a Foreign Language (EFL) learners studying at the University of Guelma - 8 Mai 1945. The selected samples embodied two groups: 30 first-year and 30 third-year students from the English Department. After assigning a writing task for students, the output was transcribed to promote the process of automated data analysis. The study employed the web-based Lexical Complexity Analyzer (LCA), introduced by Ai and Lu in 2010 for the examination of the concerned corpus. The findings showed that extensive exposure to English in academic settings does not necessarily increase the lexical density and diversity of vocabulary in students' writing. Yet, extensive exposure appears to be a prominent factor in improving the lexical sophistication of students' writings, which is considered the most important proxy measure of lexical richness. These latter findings were corroborated by the reported significant difference in the mean scores belonging to the two evaluated groups, as third-year students scored higher averages (Av. LS1= 0.83; Av. LS2= 0.69) in comparison to first-year students (Av. LS1= 0.72; Av. LS2= 0.59). The discrepancy in the mean values was statistically significant in relation to LS2 ($p = .03$) with a moderate to large effect size (LS1, $d = 0.5$; LS2, $d = 0.6$) suggesting that such a variation has practical meaningfulness, which in turn implies that the difference in the lexical sophistication of the two groups is likely noticeable to an ordinary reader. Further research in the academic context is still needed to achieve a more comprehensive understanding of the trajectory of students' vocabulary acquisition.

Keywords: artificial intelligence, natural language processing, evaluation, language proficiency, lexical richness, students' writings, lexical density, lexical sophistication, lexical diversity.

Résumé

Cette étude vise à mettre en évidence l'efficacité de l'intelligence artificielle (IA) et des outils de Traitement du Langage Naturel (TLN) dans l'analyse de textes et l'évaluation des performances des étudiants. Elle cible une dimension de la compétence linguistique, incarnée par la richesse lexicale des écrits des étudiants. Il s'agit d'une étude axée sur la mesure de la densité, de la sophistication et de la diversité du vocabulaire des étudiants. Un protocole de recherche descriptif-corrélational a été adopté, tout en utilisant une méthode quantitative d'analyse des données. Les participants étaient 60 apprenants d'anglais comme langue étrangère étudiant à l'Université de Guelma - 8 mai 1945. Les échantillons sélectionnés comprenaient deux groupes : 30 étudiants de première année et 30 étudiants de troisième année du Département d'anglais. Après avoir assigné une tâche d'écriture aux étudiants, les résultats ont été transcrits afin de favoriser le processus d'analyse automatisée des données. L'étude a utilisé l'Analyseur de Complexité Lexicale (LCA) en ligne, introduit par Ai et Lu en 2010 pour l'examen du corpus concerné. Les résultats ont montré qu'une exposition prolongée à l'anglais en milieu universitaire n'augmente pas nécessairement la densité lexicale et la diversité du vocabulaire des écrits des étudiants. Pourtant, une exposition prolongée semble être un facteur important dans l'amélioration de la sophistication lexicale des écrits des étudiants, considérée comme la mesure indirecte la plus importante de la richesse lexicale. Ces derniers résultats ont été corroborés par la différence significative signalée dans les scores moyens des deux groupes évalués, les étudiants de troisième année ayant obtenu des moyennes plus élevées (Moy. LS1 = 0,83 ; Moy. LS2 = 0,69) que les étudiants de première année (Moy. LS1 = 0,72 ; Moy. LS2 = 0,59). L'écart entre les valeurs moyennes était statistiquement significatif par rapport à LS2 ($p = 0,03$) avec une taille d'effet modérée à importante (LS1, $d = 0,5$; LS2, $d = 0,6$), ce qui suggère qu'une telle variation a une signification pratique, ce qui implique à son tour que la différence de sophistication lexicale des deux groupes est probablement perceptible par un lecteur ordinaire. Des recherches supplémentaires dans le contexte scolaire sont encore nécessaires pour parvenir à une compréhension plus complète de la trajectoire d'acquisition du vocabulaire par les étudiants.

Mots-clés: intelligence artificielle, traitement du langage naturel, évaluation, maîtrise de la langue, richesse lexicale, écrits des étudiants, densité lexicale, sophistication lexicale, diversité lexicale.

الملخص

يهدف البحث الحالي إلى تسليط الضوء على كفاءة الذكاء الاصطناعي (AI) وأدوات معالجة اللغة الطبيعية (NLP) في تحليل النصوص وتقييم أداء الطلاب. تستهدف الدراسة أحد أبعاد الكفاءة اللغوية المتجسدة في الثراء المعجمي لكتابات الطلاب. وهي تمثل دراسة تركز على قياس كثافة وتعقيد وتنوع مفردات الطلاب. تم اعتماد تصميم بحث وصفي ارتباطي مع استخدام أسلوب كمي لتحليل البيانات. شارك في الدراسة 60 من متعلمي اللغة الإنجليزية كلغة أجنبية الذين يدرسون في جامعة قلمة - 8 ماي 1945. جسدت العينات المختارة مجموعتين: 30 طالباً في السنة الأولى و30 طالباً في السنة الثالثة من قسم اللغة الإنجليزية. بعد تعيين مهمة كتابة للطلاب، تم نسخ الناتج لتعزيز عملية تحليل البيانات الآلي. استخدمت الدراسة محلل التعقيد المعجمي (LCA) المستند إلى الويب، والذي طوره Ai و Lu في عام 2010 لفحص مجموعة النصوص المعنية. أشارت النتائج إلى أن التعرض المكثف للغة الإنجليزية في البيئات الأكاديمية لا يُزيد بالضرورة من كثافة المفردات وتنوعها في كتابات الطلاب. ومع ذلك، يبدو أن التعرض المكثف عاملٌ بارزٌ في تحسين التعقيد المعجمي لكتابات الطلاب، والذي يُعتبر أهم مقياسٍ بديلٍ للثراء المعجمي. دعم هذه النتائج الأخيرة الفارق الكبير الموضح في متوسط الدرجات للمجموعتين المُقيمتين، حيث حصل طلاب السنة الثالثة على مُعدلاتٍ أعلى ($LS1 = 0.83$) ، ($LS2 = 0.69$) مقارنة بطلبة السنة الأولى ($LS1 = 0.72$; $LS2 = 0.59$). كان التباين في متوسط القيم ذو دلالة احصائية فيما يتعلق بمؤشر ($LS2, p = 0.03$) مع حجم تأثير يتراوح بين المتوسط إلى الكبير ($LS1 = 0.5$; $LS2, d = 0.6$)، مما يدل على أن للتباين دلالة عملية، و مما يعني أن الاختلاف في التعقيد المعجمي للمجموعتين يُحتمل أن يكون ملحوظاً للقارئ العادي. لا تزال هناك حاجة إلى مزيد من البحث في السياق الأكاديمي المحلي لتحقيق فهم أشمل لمسار اكتساب الطلبة للمفردات.

الكلمات المفتاحية: الذكاء الاصطناعي، معالجة اللغة الطبيعية، التقييم، إتقان اللغة، الثراء المعجمي، كتابات الطلاب، الكثافة المعجمية، التعقيد المعجمي، التنوع المعجمي.

Dedication

In the name of Allah, the most gracious, the most merciful:

{ So remember me; I will remember you, and be grateful to me and do not deny me }

All praise to Allah until it reaches its limits

To the most caring, wise, respectful and supportive man who never left my side, who taught me to place my trust in Allah, believing that what Allah has written for me is always the best, my beloved father.

To the most endlessly loving, graceful, powerful and encouraging woman I have ever known, my dear mother.

To my beautiful family and kind friends I found along the way.

Hadia

Dedication

*Above all, I thank **ALLAH** for giving me the courage and faith to carry
out this work,*

I dedicate this modest work,

*To him who has always been present, who has taught me the true values of life, who congratulates me,
rewards me for my victories, and encourages me to always go forward, my beloved **father**.*

*To the most courageous, sensitive, generous woman, most beautiful in my eyes, to the one who knew how to
give me love and joy of life, to the one who always showed affection and understanding for me, my **mother**
whom I love.*

*My dear **grandparents** for their encouragement.*

*My dear **brother** and **sister** for their trust. Whose unconditional love and constant support have been my
greatest source of strength.*

*To my **relatives** to whom I owe much gratitude for their support and prayers.*

*To all my **friends** without any exceptions.*

*To my teammate **Insaf** whom I am grateful to share this journey with.*

*To those who supported me night and day and throughout my journey. To those who helped me during my
primary studies until I graduated.*

Rania

Acknowledgment

Initially, We would like to express our praise and deepest gratitude to Allah, the most gracious, the most merciful, whose continuous guidance enlightened our path, and whose strength made this journey possible.

We would like to sincerely thank our supervisor, **Dr. Ilyes Gouider**, for his continuous guidance, valuable insights and precious pieces of advice, which were crucial to the success of this dissertation.

Our appreciation also goes to the jury members and respectful teachers, **Mrs. Yamina Biskri** and **Dr. Mahfoud Ali Zoui** who devoted their time and expertise to examine our work, besides their constructive feedback and thoughtful discussions that significantly enhanced the quality of this study.

We would also extend our gratitude to all teachers, students and educators who valued the importance of academic research and highly contributed to the successful completion of our investigation.

Table of Contents

List of Tables.....	IV
List of Figures	V
List of Abbreviations.....	VI
General Introduction	1
Statement of the Problem	1
Aims of the Study.....	2
Research Questions	3
Research Hypotheses.....	3
Significance of the Study	4
Research Methodology and Design.....	5
Research Method and Data Gathering Tools	5
Population and Sampling	5
Structure of the Dissertation.....	6
Chapter One: Literature Review.....	7
Section One: Lexical Richness in Language Proficiency	7
Introduction	7
1.1. Definition of Lexical Richness.....	7
1.2. Vocabulary Knowledge.....	9
1.3. Components of Lexical Richness	9
1.3.2. Lexical Density	10
1.3.3. Lexical Sophistication	11
1.4. The Role of Lexical Richness in Language Learning	12
1.5. Lexical Richness and Writing Proficiency.....	12
1.6. Traditional Methods for Measuring Lexical Richness	13
1.6.1. Measuring Lexical Diversity	13
1.6.2. Measuring Lexical Density	14
1.6.3. Measuring Lexical Sophistication.....	14
1.6.4. Challenges in Assessing Lexical Richness.....	16
Conclusion.....	17
Section Two: AI and NLP in Language Evaluation.....	18
Introduction	18
1.1. Definition of AI.....	19
1.2. Overview of AI and its Rapid Integration in Various Fields	20

1.3. Integration of AI in the English Classroom	22
1.4. Historical Development of NLP in Education	24
1.5. Contemporary NLP Applications in Language Learning	24
1. 6 AI-Based Tools for Evaluating Language.....	25
1.6 1. Web-Based Lexical Complexity Analyzer (LCA)	25
1.6.2. Conversational AI as a Language Learning Tool.....	25
1.7. Benefits of NLP in Education	26
1.7.1. Scalability and Efficiency	26
1.7.2. Personalized Learning and Feedback	27
1.7.3. Immediate and Continuous Assessment.....	27
1.7.4. Language Learning and Practice	27
1.7.5. Objectivity in Assessment.....	28
1.8. Ethical Considerations in the Educational Use of NLP	28
1.8.1. Bias and Fairness.....	29
1.8.2. Privacy Concerns.....	29
1.8.3. Transparency and Accountability.....	29
1.9. Practical Considerations in the Educational Use of NLP.....	30
1.9.1. Accuracy and Limitations of NLP	30
1.9.2.Over-Reliance on Automation	31
1.9.3. Cost and Accessibility	31
1.9.4. Integration into Existing Educational Practices	31
10. Future Directions for NLP in Education	32
Conclusion.....	32
Chapter Two: Research Methodology.....	33
Introduction	33
2.1. Research Design and Approach	33
2.2. Sample and Target Population	34
2.3. Research Instruments	35
2.4. Data Collection Procedures	36
2.5. Corpus Analysis	36
2.5.1. Aims of the Corpus Analysis	36
2.5.2. Corpus Selection Criteria	36
2.5.3. Description of the Corpus	37
2.5.4. Corpus Analysis Procedures.....	38
Conclusion.....	40

Chapter Three: Results & Discussion	41
Introduction	41
3.1. Descriptive Statistics	41
3.2. Inferential statistics	44
3.3. Effect Size of the Differences	48
3.4. Summary and Discussion of Corpus Analysis	49
Conclusion.....	51
General Conclusion	52
Pedagogical Implications	52
Limitations of the Research.....	53
Suggestions for Future Research.....	54
References	56
Appendix A: Writing Task.....	61
Appendix B: Samples of First-Year and Third-Year Students	62
Appendix C: Lexical Complexity Analyzer Data Processing Steps	78
Appendix D: Detailed Statistical Tables	82
Appendix E: Output from SPSS.....	84

List of Tables

Table 1. <i>Lexical Density and Sophistication Indices</i>	38
Table 2. <i>Lexical Variation and Diversity Proxy Measures</i>	39
Table 3. <i>Descriptive Statistics for Lexical Density and Lexical Sophistication Scores</i>	42
Table 4. <i>Descriptive Statistics of Lexical Variation Scores for First-year and Third-year Students</i>	43
Table 5. <i>Inferential Statistics Related to Lexical Richness Scores</i>	45
Table 6. <i>Effect Sizes of the Difference between the two Groups</i>	48
Table 7. <i>Holistic Descriptive Statistics</i>	84
Table 8. <i>Holistic Inferential Statistics</i>	85
Table 9. <i>Holistic Effect Sizes</i>	87

List of Figures

Figure 1. <i>Benefits of NLP in Education</i>	28
Figure 2. <i>Considerations for NLP in Education</i>	30
Figure 3. <i>Web-Based Lexical Complexity Analyzer</i>	78
Figure 4. <i>LCA Single Mode</i>	78
Figure 5. <i>Input Insertion into the LCA</i>	79
Figure 6. <i>Selection of Lexical Richness Indices</i>	79
Figure 7. <i>Selection of English Variety</i>	80
Figure 8. <i>Visualization and Numerical Results</i>	80
Figure 9. <i>Output from Excel Sheet</i>	82
Figure 10. <i>SPSS Dataset</i>	83

List of Abbreviations

AdjV: Adjective Variation

AdvV: Adverb Variation

AI: Artificial Intelligence

ASR: Automatic Speech Recognition

Av: Average

BNC: British National Corpus

COCA: Corpus of Contemporary American English

CVS2: Corrected Verb Sophistication II

CVV1: Corrected Verb variation-I

EFL: English as a Foreign Language

GDPR: General Data Protection Regulation

GSL: General Service List

ITS: Intelligent Tutoring Systems

L1: First Language

L2: Second Language

LCA: Lexical Complexity Analyzer

LD: Lexical Density

Log TTR: Logarithmic Type-Token Ratio

LR: Lexical Richness

LS: Lexical Sophistication

LS1: Lexical Sophistication I

LS2: Lexical Sophistication II

LV: Lexical Variation

LV: Lexical Variation

M: Mean

ModV: Modifier Variation

MSTTR: Mean Segmental Type-Token Ratio

N: Number

NDW: Number of Different Words

NLP: Natural Language Processing

NV: Noun Variation

ORC: Oxford Reference Corpus

P-value: Probability Value

SD: Standard Deviations

SLA: Second Language Acquisition

SPSS: Statistical Package for Social Sciences

SVV1: Squared Verb variation-I

TTR: Type-Token Ratio

VS1: Verb Sophistication I

VS2: Verb Sophistication II

VV1: Verb Variation-I

VV2: Verb Variation-II

General Introduction

In the digital age, the assessment of second language (L2) learners' writing development has evolved significantly, particularly in how we measure students' lexical richness across different academic levels. However, empirical evidence of this development—at least through the different stages of undergraduate study—remains limited, with traditional approaches to assessment still relying on subjective or labor-intensive analysis. The recent availability of Natural Language Processing (NLP) tools such as the Lexical Complexity Analyzer (LCA) offers new scalable methods of objectively measuring lexical features such as density, sophistication, and variation besides revolutionizing our ability to analyze and understand students' writing development. This investigation aims at explaining how academic development affects vocabulary growth over time.

Statement of the Problem

In English as a Foreign Language (EFL) context, one of the major academic challenges students face is the difficulty of retrieving the appropriate vocabulary they need to express themselves effectively. Obviously, a learner can overcome such difficulty by endorsing positive learning habits that enhance the likelihood of acquiring new lexical items. The ability to select and choose the right combination of words to convey meaning efficiently can be assessed from multiple dimensions, namely: lexical density, diversity and sophistication. All of these three fall under the broader concept of “Lexical richness”. Despite the recognized importance of vocabulary in shaping the proficiency level of learners, there has been a limited body of research exploring how lexical richness evolves across academic levels in the Algerian context. Also, very few studies attempted to gauge the level of lexical richness of students despite its feasibility in setting benchmarks for lexical competence across different educational stages.

Another research problem lies in the small number of studies that analyses students' actual performance while targeting vocabulary as an aspect of proficiency. Most of the research done so far has relied heavily on questionnaires and interviews without attempting to implement investigation tools that examine concrete data reflective of students' performance. Depending on such tools that derive information from the self-reporting of participants alone can diminish the credibility of the findings, as such methods may not accurately mirror students' true lexical proficiency. In addition, a study that attempts to empirically measure lexical richness through traditional methods can fail to meet its desired outcomes as the process of data analysis would be both time and effort-consuming.

Such methods are typically time- and labour-intensive, making large-scale empirical analysis impractical. Also, the attained findings in that case would lack reliability, since the results would be liable to be entrenched with a lot of inaccuracies. To overcome the aforementioned weaknesses, the researchers will solicit to examine lexical richness through the use of a natural language processing tool, embodied in the web-based Lexical Complexity Analyzer (LCA). Understanding the differences in lexical richness between first-year and third-year university students can provide valuable insights into the effectiveness of language instruction and curriculum design in higher education settings.

Aims of the Study

This research, on the one hand, aims at exploring how lexical richness in EFL students' academic writing develops at different academic stages at the University of Guelma. On the other hand, it focuses on analyzing variations found in lexical density, diversity, and sophistication between first-year and third-year students. By utilizing the AI tool, LCA, this study seeks also to compare the lexical proficiency of these two groups to determine any significant differences or patterns in their lexical development.

Research Questions

The present research addresses the following research questions :

- 1) How is lexical richness -measured in terms of lexical density, diversity, and sophistication?
- 2) What are the average scores of lexical richness embedded in the free writings of both groups?
- 3) How do the writings of first-year and third-year EFL students differ in terms of vocabulary use?
- 4) How do language proficiency and academic development relate to lexical richness development in higher education?

Research Hypotheses

The research hypotheses have been formulated to examine the differences in lexical richness between first-year and third-year EFL students' written output:

1-. H_0 (null): There is no significant difference in the mean scores of lexical density reflected in the written output of first-year and third-year EFL students.

- H_1 (alternative): There is a significant difference in the mean scores of lexical density reflected in the written output of first-year and third-year EFL students.

2- H_0 (null): There is no significant difference in the mean scores of lexical sophistication.

- H_1 (alternative): There is a significant difference in the mean scores of lexical sophistication.

3- H_0 (null): There is no significant difference in the mean scores of lexical diversity and variation.

$-H_1$ (alternative): There is a significant difference in the mean scores of lexical diversity and variation.

Significance of the Study

The significance of this study lies in filling the research gap related to investigating the lexical richness of students' free writings in the Algerian context of education. This dissertation adheres to the standards of scientific research as the conduction of the study went through the phases that underlie empirical inquiry. The study extracts data from a corpus comprised of texts written in real-time, which are then examined through the use of computational data analytic tools.

Moreover, the study enriches the existing literature on lexical complexity development in higher education and the obtained findings can help in establishing benchmarks for the normative scores belonging to the two academic levels. Thus, language educators can use the findings as a reference point in assessing the proficiency level of first-year and third-year students. The findings of this research can inform language educators, curriculum developers, and policymakers about the importance of fostering lexical growth of university students by integrating AI into the evaluation process, particularly in the assessment of students' lexical richness more efficiently, more accurately, timelessly and effortlessly. The statistical data are thoroughly illustrated and the incorporated method is elaborated in detail to promote the replicability of the present design. The application of the design in future studies by other researchers can help in forming generalisations about the phenomena under investigation.

Research Methodology and Design

Research Method and Data Gathering Tools

This research employs a descriptive-correlational research design in answering the research questions and testing hypotheses. A corpus analysis is implemented as a tool to collect data from both first-year and third-year students, at the Department of English, 08 May 1945 University, Guelma. The corpus consists of sixty (60) written texts produced by students of both groups. These texts serve as the basis for analyzing lexical richness features of the participants' written texts to measure the density, diversity, and sophistication of their vocabulary.

Population and Sampling

The study targets first-year and third-year undergraduate students who belong to the English Department at the University of 08 May 1945, Guelma. A total sample of sixty (60) students was selected; thirty (30) from each academic year. This type of selection is known as purposive sampling since those students who were involved in the data collection procedure met the demographic criteria of the research and served the objective set at the outset of the study. It can be also considered a convenience sampling method from the perspective that the present researchers worked on samples of students who were available or accessible during the data collection period. The aspect of availability relates to the willingness of students to take part in the study, whereas accessibility pertains to the teachers' consent in granting access to the concerned classes.

The selection of these two groups enabled the conduction of comparative corpus-based analysis between intermediate and upper intermediate/advanced learners to better understand the progression of their lexical development.

Structure of the Dissertation

The current dissertation is organized into 3 chapters, along with a general introduction and a general conclusion. The introductory section gives an overview of the nature of the research alongside its rationale, problem statements, research gap and significance. The first chapter involves a literature review composed of three sections. The first one, *Lexical Richness in Language Proficiency*, provides a comprehensive theoretical foundation for the research. It investigates lexical richness covering its definition, main components (lexical variety, lexical density and lexical sophistication), role in language learning, measurement methods as well as the potential challenges faced in assessment. The second section, *AI and NLP in Language Evaluation*, presents an overview of the integration of Artificial Intelligence and Natural Language Processing tools, such as the Lexical Complexity Analyzer (LCA), in the context of education.

The second chapter elaborates on the methodology of research in detail as it describes the research site, the population and the sample and explains the research design, the procedures and instruments involved in data collection and analysis, as well as the statistical methods employed for the obtainment of the results. The third chapter presents the results through the use of descriptive and inferential statistics supplemented by visual representations such as tables, pie charts and bar charts. The chapter also discusses the results and provides the perceived implications of the findings accompanied by recommendations. The last section involves a general conclusion that identifies the insights obtained through the process of data analysis and interpretation. The encountered limitations are also mentioned along with the suggestions for future research.

Chapter One: Literature Review

Section One: Lexical Richness in Language Proficiency

Introduction

Language proficiency, particularly in the context of foreign language learning, is commonly gauged through the learner's ability to effectively communicate across different contexts and settings. The concept of lexical richness is considered a cornerstone element of evaluation that reflects the vocabulary knowledge possessed by learners. Due to the fact that educators as well as researchers focus on enhancing language acquisition, the role of lexical richness emerged as a significant indicator of language proficiency, influencing both spoken and written discourses.

Lexical richness is a central feature of language use. It is believed that good writing is characterised by these lexical features that constitute lexical richness: lexical variation, lexical density, and lexical sophistication in addition to a low incidence, if not total absence, of lexical errors (Read, 2000). In research areas such as stylistics, language assessment, text readability analysis and first language acquisition, vocabulary richness is regarded as a core element that qualifies the depth, complexity, and overall quality of language use (Lu, 2012). This chapter solicits to explore the role of lexical richness in language proficiency by providing its definition, highlighting its role in language learning, examining its major components (lexical diversity, lexical density and lexical sophistication), identifying its role in determining writing quality, discussing common methods used to gauge these aspects, while considering the potential challenges which may occur when measuring lexical richness.

1.1. Definition of Lexical Richness

Lexical richness is a multidimensional concept, as it is an essential component of language proficiency, serving as an indicator that specifically assesses the quality of

language production in both oral and written discourses of EFL learners. It reflects the learners' ability to effectively use a diverse, dense, sophisticated and complex vocabulary in their expression (Read, 2000). It provides insight into a learner's linguistic development and communicative competence. This concept is quite important for understanding both the depth and breadth of a learner's vocabulary knowledge, as well as the capacity to apply it in different real-world situations.

In consonance with Yule (1944), richness of the vocabulary is defined as the wealth of an author's vocabulary. Sometimes, when reading a work, one may feel a strong impression that the writer possesses a remarkable wealth of words, signifying a rich vocabulary based on this subject judgement. Building on this definition, Yule is credited as one of the pioneering scholars to establish a clear framework for the concept of lexical richness. Lu (2012) proposed that lexical richness is manifested in the sophistication and range of a learner's productive vocabulary. In simpler terms, he related lexical richness to the degree of advancement and variation in the vocabulary a learner utilizes in written and spoken discourses, thus addressing it from a different angle.

At its core, lexical richness is composed of several fundamental components that work to provide a holistic view of vocabulary usage. These components include lexical density, which examines the proportion of content words to the total number of words in a text; lexical diversity, which assesses the variety of different words employed; and lexical sophistication, which considers the use of less frequent or more advanced vocabulary items (Erando & Fortes, 2024). The conceptualization of lexical richness has evolved over time. Initially, it was used in first language (L1) contexts and literary stylistics, to demonstrate the variety, complexity, sophistication and density of vocabulary used in spoken or written language. However, the term lacked a clear meaning and methodological context,

especially in the field of second language acquisition (SLA), despite being often used in academic literature.

1.2. Vocabulary Knowledge

Vocabulary knowledge refers to how well words are known, including their meaning, form and appropriate use in different settings. Meaning points to relating the word to its actual sense, concept, and referent, as well as its association with other words. Whereas form refers to knowing the spoken and written form of a word and being able to easily identify its parts. More precisely, this construct is based on the notion of knowing the grammatical functions, and collocations, alongside acknowledging the limitations of the usage of that word (Nation, 2013; Read, 2000, as cited in Kilic, 2019). Equally important, vocabulary knowledge is not only about vocabulary size, i.e. knowing a large number of words, but rather about how well we know and understand words (vocabulary depth), how many words we know (vocabulary breadth), and being capable of using those words appropriately across different language skills, whether in receptive or productive skills (McCarthy, 2000; Mehrpour & Rahimi, 2010; Ouellette, 2006; Qian, 1999; Shiotsu & Weir, 2007, as cited in Kilic, 2019).

1.3. Components of Lexical Richness

Lexical richness mainly consists of three key components, lexical diversity, lexical density and lexical sophistication, each of which provides a different lens through which the vocabulary richness of a learner can be assessed.

1.3.1. Lexical Diversity

A key component in lexical richness is lexical diversity (also called lexical variety). It is defined as the variety of the variety of lexical items employed in a text, indicating the ability of a learner to effectively employ a wide range of vocabulary. In the context of

learning EFL, lexical variety serves as an indicator of language proficiency as well as a learner's capacity to articulate ideas using their acquired lexicon. Lexical diversity plays a significant role in both receptive and productive language skills, contributing to a learner's overall linguistic competence (Nation, 2001). Its development is influenced by several factors including exposure to language input, language learning strategies, and individual differences (Kalantari & Gholami, 2017). Kyle (2019) asserted that this measure is particularly important in language learning, as higher proficiency language learners are expected to produce language that includes a wider variety of lexical items, demonstrating a more extensive vocabulary range.

The traditional way for assessing Lexical variation is through calculating the type-token ratio (TTR), by dividing the number of unique words (types) by the total number of words (tokens) in a text. While TTR offers a straightforward computational method, it is inherently sensitive to text length; as the length of a text increases, the likelihood of word repetition rises, thereby reducing the TTR value and complicating cross-textual comparisons (Jarvis, 2002; McCarthy & Jarvis, 2007; Read, 2000).

1.3.2. Lexical Density

Lexical density is another valuable measure of lexical richness, which refers to the proportion of lexical words, also known as content words (including nouns, verbs, adjectives, and adverbs), to the total number of words in a text. Besides, it may also be used to indicate students' progress in learning a language, particularly in writing (Syarif & Putri, 2018), enabling educators to evaluate both the efficiency of teaching materials and students' writing abilities.

In other words, lexical density is the proportion of words in written or spoken language that give us an idea about what is being communicated. This implies that lexical

density serves as an indicator of text's informativity (Analyze My Writing, n.d.), demonstrating how much information is conveyed through the content words. Johansson (2008) wrote that according to Ure's analysis of measuring lexical density, spoken discourse has a lexical density below 40%, while written discourses are characterized by a lexical density of 40% or higher. These findings were obtained from 34 spoken text and 30 written text samples, as well as manually calculated. That is to say, written language tends to be more rich in vocabulary than spoken language.

Lexical density serves as a crucial indicator of EFL learners' language proficiency. A study conducted on Saudi EFL undergraduate students discloses that lexical density increases as students progress to higher levels of education, as it also varies according to learner proficiency because of the obtained differences in lexical density scores between samples (Elgobshawy & Aldawsari, 2022). Accordingly, lexical density presents an important index of a learner's linguistic development.

1.3.3. Lexical Sophistication

Lexical sophistication, also referred to as "Rareness", is defined as the use of sophisticated words in a learner's text, as it embodies another significant component of lexical richness. Read (2000) in his book "Assessing Vocabulary Knowledge" affirmed that lexical rareness refers to those advanced words that learners are not expected to know compared to their educational level. Accordingly, the use of low-frequency words is a strong indicator of higher lexical sophistication. A study conducted by Ha (2019) on 35 EFL undergraduate students at Korea University in Seoul found lexical sophistication to be the most influential factor that determines the high writing quality of a learner, which is directly associated with higher language proficiency, demonstrating a higher lexical richness.

1.4. The Role of Lexical Richness in Language Learning

In an EFL context, lexical richness occupies a central role in determining language proficiency and the quality of academic writing.

Overall writing quality may indicate that students with abroad vocabulary dictionary have a higher level of English language proficiency (Kim, 2014; Lemmouh, 2008) [...] The lexical richness displayed in written text is a result of a person's underlying vocabulary knowledge, which can be effective in academic writing. (Ha, 2019, pp. 21-23)

This suggests that a rich vocabulary enhances writing quality, indicating a higher proficiency in English, especially in academic contexts where varied and precise language is exclusively important. Moreover, it also emphasizes the notion that the level of lexical richness in writing reflects a learner's vocabulary knowledge, which is foundational for language learning, therefore, the more words learners acquire, the richer their language becomes.

1.5. Lexical Richness and Writing Proficiency

Vocabulary richness falls under the umbrella of vocabulary knowledge, which contains both receptive and productive skills of language, particularly through lexical sophistication, LR (lexical richness) holds a crucial role in determining writing quality and proficiency (Read, 2000). The findings of a study conducted by a Tunisian researcher, Ayadi (2023), have shown that indices like lexical rareness of GSL-1 1000 words strongly correlate with higher writing quality, whereas lexical density and lexical variation were not statistically significant. This latter highlights the need for carefully considering lexical sophistication when developing lesson plans that address writing skills. Additionally, measuring lexical richness helps teachers and educators to generate an idea about their students' writing abilities as well as to assess their flaws and weaknesses.

1.6. Traditional Methods for Measuring Lexical Richness

1.6.1. *Measuring Lexical Diversity*

The traditional manual measurement of lexical diversity is the Type-Token Ratio (TTR), calculated through dividing the number of different words (types) by the total number of words (tokens), i.e.:

$$\text{TTR} = \frac{\text{the total number of different words}}{\text{the total number of words}}$$

However, if the text sample contains a large number of tokens, the TTR values will lower and vice versa (Johansson, 2008). Johansson (2008) also suggested that: “TTR is only possible to use when comparing texts of equal length. In spite of this, TTR is still used for comparing text production” (p. 63). We can conclude that despite the utility of such measurement in determining how varied a student’s language is, it has limitations. The drawback arises when text length increases, as the TTR value drops, making it challenging to compare students’ text samples that vary in length.

VocD or D measure, developed by David Malvern and Brian Richards in 1997 (MacWhinney, 2000; Malvern et al., 2004; Malvern & Richards, 1997, as cited in Johansson, 2008) is another critical index for evaluating variance that is independent of sample size (length), solving the issue of TTR by allowing the comparison of texts that are different in length. However, it received some criticism from scholars in the field for not serving this goal. In addition to VocD, a third common indicator of lexical variety is the so-called Guiraud Index, which was developed to fulfil the same aim as the D value, i.e., to be able to deal with texts that have unequal size (Johansson, 2008), working as an alternative to compensate for some of the limitations of the TTR.

1.6.2. Measuring Lexical Density

From a statistical and quantitative view, Ure (1971) suggested a slightly different method to measure lexical density, by calculating it as the percentage of lexical or content words to the total number of words in a text, i.e.:

$$\text{Lexical density (\%)} = (\text{Number of lexical words} / \text{Total number of tokens})$$

Halliday (1985) introduced an approach referred to as “Halliday’s Method” for calculating lexical density, which is as follows:

$$\text{Lexical density} = \text{Total number of lexical items} / \text{Total number of clauses}$$

Therefore, he offered a measurement from the perspective of functional grammar, where lexical items or lexical words are also labeled as content words. For example, consider the sentence: “Sunny spent the whole morning preparing fluffy pancakes, while her cat watched curiously from the window”. The total number of its lexical words is (12), and the total number of its clauses is (2). By applying Halliday’s formula to the context, we conclude that:

$$\text{The lexical density of the sentence} = \text{Content words} / \text{Total N of clauses} = 12 \div 2 = 6$$

In addition to these traditional manual methods, nowadays several advanced computational tools have been developed to measure lexical density, such as web-based Lexical Complexity Analyzer (LCA), Writing Style Analysis Tool, available at ReadabilityFormulas.com, Analyze My Writing – Lexical Density Calculator, among other tools.

1.6.3. Measuring Lexical Sophistication

There are several methods to gauge lexical rareness in a given language sample. Using word frequency lists derived from a reference corpus, in order to determine how

sophisticated a word is, is one of the well-attested ways of examining words' frequency (Gouider, 2023 a; Kyle & Crossley, 2014). This method works simply by comparing the words of a text sample to the words found in such lists, thereby determining their degree of frequency.

The primary goal for measuring lexical sophistication is to have an idea about how advanced or low-frequent the vocabulary in a text is, where the use of more advanced words indicates a higher level of writing or speaking proficiency. One of the traditional methods which were used earlier, before computational tools came to exist, is the GSL or General Service List. GSL, created by Michael West in 1953, is a list of 2000 words that are commonly or frequently used (such as “home”, “the” and “possible”), that appear across a wide range in English texts (Smith, n.d.). Therefore, GSL was used as an indicator of high-frequency or low-sophisticated vocabulary. In light of this, to measure lexical sophistication in a given text, teachers used to compare vocabulary incorporated in a text sample to the words in such standardized reference corpus of English, where words that are not included in the list would be classified as low-frequency words, therefore considered highly sophisticated and vice versa. However, the GSL has been developed and refined over time to include updated words, and similar corpora, such as Oxford Reference Corpus (ORC), British National Corpus (BNC) and Corpus of Contemporary American English (COCA), to reflect more contemporary language usage.

In fact, while such traditional manual methods for assessing lexical richness (lexical variety, lexical density and lexical sophistication) were once essential in language analysis, they were not been replaced by today's computational tools, rather, they actually led to their existence. Modern tools such as the web-based Lexical Complexity Analyzer, Coh-Metrix and TAALES still rely on the exact principle calculations established by

earlier methods. The only difference lies in the automation, scalability, efficiency and accuracy offered by such methods.

1.6.4. Challenges in Assessing Lexical Richness

Measuring lexical richness may present several challenges that researchers must navigate to ensure accuracy in the assessment of language, due to the sensitivity of defining as well as quantifying its components. A primary key issue is related to measuring lexical diversity, precisely to the TTR index, due to the influence of text length, resulting in a greater number of tokens, on the number of types a text contains (Bestgen, 2024). Making comparisons across text samples of different sizes is problematic as mentioned earlier. Hence, this directly leads to reducing the reliability and validity of the index. From another angle, Zhang & Wu (2021) put forward the following:

A clear-cut definition is still missing: the term 'lexical proficiency' refers to vocabulary size, the depth of word knowledge, and the degree of sophistication of word use (Crossley et al., 2010). A survey of existing studies, where lexical proficiency is measured quantitatively, shows that lexical richness (LR) is most widely used, both conceptually and practically, for research along the lines of the current study, although the measurement of LR has proven quite an open problem in its own right (Jarvis, 2013; Malvern et al., 2004; Tweedie & Baayen, 1998). (p. 2)

As evidenced above in the quotation, a major limitation an educator can face when assessing lexical richness is the absence of a clear widely agreed-on definition of lexical proficiency. Furthermore, Yanhui Zhang and Weiping Wu highlighted that despite the extensive use of lexical richness in both theoretical and applied linguistic research, its measurement continues to present a challenge.

The absence of a standardized approach to evaluate lexical richness, as well as the existence of different calculation methods for the same vocabulary richness measurement

are other notable shortcomings. Taking lexical density as an example, across time, many researchers provided varied definitions with varied measurements of the former. For instance, as noted earlier, Halliday (1985) and Ure (1971) have proposed divergent interpretations and methodologies, each one approaching lexical density from distinct viewpoints.

Another separate difficulty occurs when evaluating LR (lexical richness) following computational approaches and programs. Due to the fact that such mechanized tools tend to focus on the form such as language length and word frequency, disregarding the meaning conveyed through that language. That is to say, they may prioritize the surface structure of a student's texts over its deep structure, which may lead the educator to make a wrong judgement about the sample.

Conclusion

In closing, LR serves as an important index of student's learning proficiency, exclusively in the context of English as a Foreign Language (EFL) education, reflecting the breadth of vocabulary knowledge, also the ability to employ it effectively across varied settings. This chapter has investigated extensively the concept of lexical richness. It also illustrates the main components involved in its evaluation (lexical diversity, lexical density and lexical sophistication) which are exclusively central for understanding the complexity of language use. Hence, these fundamental elements contribute to assessing learners' vocabulary knowledge and, consequently, their overall linguistic competence, more precisely writing and speaking. Nevertheless, despite that a lot of limitations in evaluating vocabulary richness continue to exist, both manual and automated tools for assessing it remain essential to ensure more accurate, validated, scalable as well as reliable measures of lexical richness.

Section Two: AI and NIP in Language Evaluation

Introduction

Artificial intelligence (AI) has received a lot of attention recently as a possible tool to innovate teaching and learning in educational environments. However, the concept of using AI in education is not new. Since Turing (1950) explored the mathematical possibility of AI by proposing the Turing Machines, the progress of AI and research on learning and educational applications had ups and downs (as cited in Ji, Han & Ko, 2022). Minsky and Papert (1968) pioneered AI research and applied computational theories to imitate human psychological processes in machines so that a computer will be able to solve issues and make decisions based on available knowledge in the same way as humans do (as cited in Ji, Han & Ko, 2022). Since then, early research on education and AI concentrated on the development and testing of intelligent tutoring systems (ITS). Combined with the pedagogical approach of learning by teaching, some ITSs have evolved into teachable agents that assist students learn while teaching computer agents (Blair et al., 2007; Leelawong & Biswas, 2008; Silvervarg et al., 2021, as cited in Ji, Han & Ko, 2022).

AI has gained new momentum with the recent increase in investment in the industry (Pan, 2016, as cited in Ji, Han & Ko, 2022). It has accelerated the development of AI technology, including neural networks, machine learning, natural language processing (NLP) with automatic speech recognition (ASR), and advanced image processing (Zawacki-Richter et al., 2019, as cited in Ji, Han & Ko, 2022). It has also increased access to commercially available AI devices and mobile applications. Easier and wider access to AI technologies enabled a potential growth in its application in educational contexts. This was corroborated by a considerable increase in the number of papers published on this topic since 2015 (Chen et al., 2020, as cited in Ji, Han & Ko, 2022). Indeed, a review by Chen et al. (2020, as cited in Ji, Han & Ko, 2022) showed a wide range of educational

settings, where modern AI technologies were used in instruction, administration , and learning. The settings included tailored intelligent teaching, assessment and evaluation, smart schools, and remote education via online and mobile devices .

1.1. Definition of AI

Artificial Intelligence (AI) is a field of computer science developed to replicate the activity and performance of the human brain through algorithms. It has the ability to learn by example and simulation, creating predictions based on the knowledge it collects (Abiodun et al., 2018, as cited in Hider ,2024). Hider (2024) argued that in order to define Artificial Intelligence, one must first look back to the origins of the term. Thanks to John McCarthy, often referred to as the "father of AI", and the one who originated the term "artificial intelligence". He added that artificial intelligence is the field that focuses on producing machine intelligence rather than replicating the human intelligence, it does not only mimics human intelligence but also it learns it. This later allows the software to find solutions and perform tasks on its own without relying on of human intelligence stimulation through solutions that may not exist in the parallel human world; thus, this comprehension exceeds the constraints that humans are biologically bounded by. Though the two definitions demonstrate some inconsistencies, AI indeed is still in the stimulation phase.

According to Hider (2024), Encyclopedia Britannica differently defines Artificial Intelligence as the ability of a digital computer to execute tasks typically associated with intelligent beings. Similarly, during a virtual scientific seminar, Alcina (2008) defined AI as the intelligent use of data to assist humans emphasizing that the goal is to produce software that can think, sense, act, manage and adapt, enabling machines to think through data and algorithms (as cited in Hider ,2024)

Despite the different views in defining artificial intelligence, it is agreed that AI is the stimulation of human brain intelligence in performing tasks via software that utilizes algorithms to complete tasks.

1.2.Overview of AI and its Rapid Integration in Various Fields

The integration of artificial intelligence (AI) has rapidly spread across numerous fields, presenting both challenges and opportunities. It has the ability to increase efficiency, production, and service delivery. However, it also necessitates careful consideration of ethical concerns, the creation of suitable policies and regulations, and ongoing involvement of human oversight. Addressing these problems allows AI to effectively integrated into numerous sectors, contributing to developments and improvements in different domains.

Self-learning algorithms, a core component of AI, have created new opportunities and had a significant impact on various sectors. AI has gained importance and has been recognized in the public sector for its ability to create new opportunities (Wirtz et al., 2018, as cited in Özdere, 2023). AI in the public sector can optimize organizational structures, improve work productivity, and resolve management problems (Lu & Gao, 2022, as cited in Özdere, 2023). It has the ability to enhance the services given by governmental organizations, making them more interactive and user-friendly (Almaiah et al., 2022, as cited in Özdere, 2023). AI has been utilized in market and public administration to optimize organizational structures, process data, and enhance work efficiency (Lu & Gao, 2022, as cited in Özdere, 2023). The use of AI in recruitment and selection processes can impact applicants' perceptions and reactions, depending on the design features of AI assessments and the positioning of AI tools in the hiring process (Hunkenschroer & Luetge, 2022, as cited in Özdere, 2023). However, the use of AI in the public sector necessitates careful consideration of variables such as trialability, observability,

complexity and compatibility. Understanding the importance of these factors and planning accordingly can help effectively implement AI in governmental systems (Almaiah et al., 2022, as cited in Özdere, 2023). Concerns concerning about the impact of AI on the employment market , as well as potential biases and inequalities in AI systems must be addressed (Qadir, 2022, as cited in Özdere, 2023).

The use of artificial intelligence (AI) in education is a rapidly expanding field with several applications and ramifications. AI technologies, powered by machine learning algorithms, have the ability to transform teaching and learning processes, improve educational outcomes, and enhance the whole educational experience. AI has been applied to various aspects, including language education, feedback analysis and literature analysis. AI-powered tools can boost language learning experiences, analyze student feedback, and improve educational infrastructure and teaching procedures (Raj, 2023; Shaik, 2023, as cited in Özdere, 2023).

AI technology can also be used in higher education to assist with teaching and learning processes. AI-powered systems can assess student data for tailored personalized recommendations, adaptive learning, and increase student engagement (Kashive et al., 2020, as cited in Özdere, 2023). Additionally, AI can automate administrative work like grading and scheduling, allowing instructors to focus more on instructional activities (Kashive et al., 2020, as cited in Özdere, 2023).

However, integrating AI in education has challenges such as teachers' lack of expertise about AI technology and ethical concerns (Shaik, 2023; Liao, J., Y., 2022, as cited in Özdere, 2023). Teachers and researchers must ensure that AI technologies are implemented ethically and responsibly in educational settings (Dahmash et al., 2020). To

effectively integrate AI into the curriculum , teacher training and professional development is needed (Wood et al., 2021, as cited in Özdere, 2023).

Overall, AI has rapidly merged into many industries, presenting both potential and challenges. It has the ability to increase productivity, optimize operations, and help accomplish sustainable development goals. However, ethical concerns, the need for teacher preparation, and the possible influence on inequality and transparency must all be addressed in order to ensure the appropriate and effective integration of AI across sectors.

1.3. Integration of AI in the English Classroom

AI integration in the English classroom refers to the use of artificial intelligence technology and tools into language teaching and learning processes. According to Özdere (2023) AI-powered language learning platforms, whether online platforms or mobile applications, provide engaging and adaptable learning experiences. These platforms use AI algorithms to analyze learners' performance, provide individualized feedback, and offer tailored language training. They may include language activities, vocabulary drills, grammar explanations, and interactive simulations to help students develop their language skills.

Incorporating AI into the English classroom yield good results in various studies. Li and Peng (2022, as cited in Özdere, 2023) found that the integration of an AI-based language learning platform with a flipped classroom instructional paradigm, resulted in positive attitudes and enhanced engagement among students. Zhao and Nazir (2022) apply AI and online reading platforms to improve English multimode production and usage, enabling personalized and immersive language learning experiences(as cited in Özdere, 2023). Fitria (2021, as cited in Özdere, 2023) investigates the impact of AI in the EFL environment, stressing the development of intelligent teaching systems, smart classrooms, and AI-based English teaching aids .Alhalangy (2023, as cited in Özdere, 2023)

demonstrates the efficiency and reliability of automated scoring systems, focusing on the automation of video assessment in instructional learning. Meldia and Zakir (2022, as cited in Özdere, 2023) address the problems faced by educators in integrating ICT into English teaching and learning, emphasizing the necessity of training, competencies, and technical abilities.

AI has been applied into language training to enable personalized and adaptable learning experiences. Adaptive learning platforms utilize AI algorithms to monitor learners' performance, identify their strengths and weaknesses, and provide individualized learning materials and activities. Natural language processing algorithms may evaluate and extract information from vast amounts of text, allowing for automated text summarization, language understanding and sentiment analysis (Huang et al., 2021, as cited in Özdere, 2023). The AI technologies analyze and evaluate learners' language performance, providing individualized feedback and assistance using natural language processing and machine learning algorithms (Fu et al., 2020, as cited in Özdere, 2023). They can help language learners understand and analyze difficult texts, boosting their reading and comprehension abilities. In addition, AI-powered writing tools can improve writing skills of students by offering tailored feedback and guidance. It was argued that using AI powered writing tools can increase students' behavioral, emotional and cognitive engagement, and self-efficacy for writing (Nazari et al., 2021, as cited in Özdere, 2023).

These studies demonstrate the possibility of AI integration in the English classroom, including the use of AI-based platforms, chatbots, automated assessment, multimodal education, and intelligent teaching systems. AI technologies enable tailored learning, increased engagement, effective evaluation, and immersive language experiences. However, difficulties including training, ethical considerations, technical skills, and

effective integration must be addressed in order to maximize the benefits of AI in the English classroom.

1.4. Historical Development of Natural Language Processing (NLP) in Education

The integration of Natural Language Processing (NLP) in education started with simple tools for improving text processing and analysis. Early implementations included grammar checkers and spelling which used rule-based ways to improve writing accuracy. As machine learning techniques improved, these systems evolved to include statistical methodologies and more sophisticated algorithms for better language understanding and evaluation (Kukich, 2000, as cited in Kumar & Howard, 2024). The transition from rule-based to data-driven models was a big step forward, enabling more complex language processing and paving the door for automated feedback systems.

1.5. Contemporary NLP Applications in Language Learning

Recent advances in NLP have resulted in the development of a number of sophisticated language education tools. Automated essay scoring systems like the Educational Testing Service's e-rater, have been critical to this shift. These systems use NLP techniques to evaluate writing quality based on syntax, content relevance and coherence (Attali & Burstein, 2006, as cited in Kumar & Howard, 2024). Similarly, services like Grammarly use NLP to deliver real-time grammar and style corrections, increasing the writing skills of users by making contextually relevant suggestions (Gonzalez & Smith, 2020, as cited in Kumar & Howard, 2024).

Conversational agents and chatbots are another important application of NLP in language acquisition. Tools like Duolingo and Rosetta Stone use NLP to evaluate spoken and written responses, providing immediate feedback and individualized learning ways. These systems imitate interactive language practice environments by combining machine

learning models and natural language understanding (Huang & Zhao, 2021, as cited in Kumar & Howard, 2024).

1. 6 AI-Based Tools for Evaluating Language

1.6 1. Web-Based Lexical Complexity Analyzer (LCA)

Ai and Lu (2010) developed the Linguistic Complexity Analyzer (LCA), a tool for carefully assessing several linguistic complexity characteristics within a given text sample. In order to give a full study of the complexity levels available in the language data, components such as lexical diversity, syntactic structure, morphological variation, and discourse aspects were considered. The Lexical Complexity Analyzer (LCA), established in 2010 by Haiyang Ai and Xiaofei Lu. It is an electronic tool for analyzing the lexical complexity of English texts. by drawing on earlier linguistic development studies of first and second languages. Drawing on previous studies of first and second language development, it measures 25 different indices related to vocabulary density, variety, and sophistication.

1.6.2. Conversational AI as a Language Learning Tool

Recent improvements in machine learning, ASR, and NLP technology have made conversational AI-integrated language learning a more appropriate and cost-effective strategy. They give language learners access to language learning resources and a genuine setting for communicating in a target language.

Furthermore, conversational AI can minimize foreign language anxiety, a persistent issue that impedes language learning performance and achievement (Horwitz, 2001; Shao et al., 2019; Teimouri et al., 2019, as cited in Ji, Han & Ko, 2022). Shao et al. (2019, as cited in Ji, Han & Ko, 2022) stressed the importance of positive emotions like pride, hope and contentment in language learners' motivation and performance. It was also suggested that teachers should work to lesson students' foreign language anxiety in and out of the

classroom. Furthermore, language learners have more flexibility when interacting with conversational AI (Zhang & Zou, 2020, as cited in Ji, Han & Ko, 2022). They can also receive scaffolding and criticism in a less intimidating environment than in typical classrooms (Bibauw et al., 2019; Istrate, 2018, as cited in Ji, Han & Ko, 2022). Conversational AI serves as both a pedagogical tool and an evaluative framework within modern language education.

1.7. Benefits of NLP in Education

NLP technologies provide significant benefits to the educational industry, particularly in improving learning experiences and outcomes for students:

1. It provides personalized learning.
2. Enhances automation and adaptability.
3. Promotes in language learning.
4. Bridges gaps.
5. Automated assignment scoring.
6. Ensures efficient use of resources.

1.7.1. Scalability and Efficiency

One of the most significant benefits of NLP in education is the capacity to scale evaluation and feedback systems. Automated systems can grade essays, analyze written responses, and deliver feedback to a large number of students in a timely and effective manner. This scalability is especially useful in big class sizes, when teachers may struggle to deliver tailored feedback to each student. Automated essay scoring methods, for example, can help reduce the burden on teachers, allowing more tailored instruction (Attali & Burstein, 2006, as cited in Kumar & Howard, 2024).

1.7.2. Personalized Learning and Feedback

NLP systems can assess student writing and language use, providing real-time feedback suited to individual learning requirements. This individualized feedback enables students to enhance their language skills by identifying particular areas for development, such as grammar, sentence structure, or vocabulary usage. Unlike traditional assessment approaches that provide broad feedback, NLP systems can provide specific suggestions, improving the learning experience (Gonzalez & Smith, 2020, as cited in Kumar & Howard, 2024). This level of personalization enables students to learn at their own pace while receiving immediate feedback, encouraging continuous development.

1.7.3. Immediate and Continuous Assessment

Traditional evaluations, including exams and essays, frequently deliver feedback only after a significant wait. However, NLP-based systems can provide fast feedback, allowing students to correct errors and learn in real-time. This continuous evaluation technique helps students to immediately recognize and correct mistakes, resulting in a more dynamic learning experience (Wang & Liu, 2021, as cited in Kumar & Howard, 2024). It also encourages a formative approach to learning, in which feedback is used to foster progress rather than merely an assessment of prior achievement.

1.7.4. Language Learning and Practice

NLP-powered conversational agents, chatbots, and virtual teachers allow students to practice language in a relaxed setting. These systems imitate real-world conversations, allowing students develop their speaking and writing abilities. Language learning programs such as Duolingo, use NLP to evaluate student replies and provide feedback on pronunciation, grammar, and vocabulary usage (Huang & Zhao, 2021, as cited in Kumar & Howard, 2024). This interactive and engaging approach to learning can boost student motivation and help them improve their language skills.

1.7.5. Objectivity in Assessment

Human grading is subjective, with evaluations potentially influenced by factors such as exhaustion or unconscious bias. NLP systems provide a more consistent and objective approach to grading, particularly in tasks such as essay scoring, where characteristics like grammar, content relevance and coherence may be examined using standardized approaches (Attali & Burstein, 2006, as cited in Kumar & Howard, 2024). By reducing human bias, NLP tools help to make evaluations more fair.

Figure 1

Benefits of NLP in Education



1.8. Ethical Considerations in the Educational Use of NLP

The growing use of Natural Language Processing (NLP) in educational contexts has highlighted various ethical considerations that must be properly addressed. Kumar and Howard (2024) claimed that these challenges include bias, fairness, privacy, transparency, and the potential over-reliance on automated systems.

1.8.1. Bias and Fairness

One of the most pressing ethical concerns in NLP applications is the prevalence of bias in the training data. NLP models use large datasets to learn language patterns, but if these datasets contain biased information, the programs may unintentionally perpetuate these biases. This can result in unjust evaluations of students, especially those from underrepresented or underprivileged groups. For example, an NLP-based automated essay grading system may evaluate students unfairly due to slight linguistic differences reflecting cultural or socio-economic backgrounds (Binns, 2018, as cited in Kumar & Howard, 2024). Ensuring that training data is diverse and representative is critical for reducing prejudice and increasing fairness in educational assessments.

1.8.2. Privacy Concerns

As NLP systems frequently rely on student data to individualize comments and deliver personalized evaluations, issues about data privacy and security arise. Schools and educational institutions must protect sensitive student data and keep students informed about how their data is being used. In an age where data misuse of personal information is becoming increasingly prevalent, secure student data should be a top priority. This necessitates explicit procedures for data storage, usage, and sharing, as well as adherence with relevant data protection regulations, such as the General Data Protection Regulation (GDPR) (Tsamados et al., 2022, as cited in Kumar & Howard, 2024).

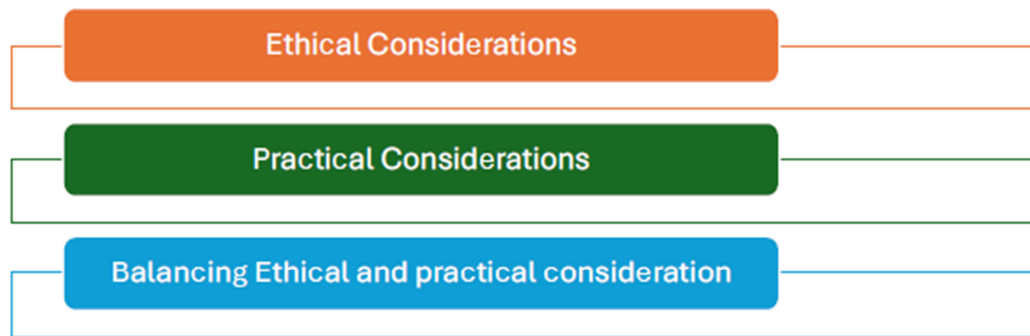
1.8.3. Transparency and Accountability

Another ethical concern is the lack of transparency in how NLP models make choices. In many circumstances, NLP systems are viewed as "black boxes," where it is difficult to grasp the decision-making processes. This makes it difficult to explain to students and teachers why a specific assessment or feedback was generated. Ensuring transparency and accountability in NLP systems is critical to building confidence.

Educators and students should understand how these systems operate and have avenues for recourse in the event of inaccurate assessments (Selbst et al., 2019, as cited in Kumar & Howard, 2024).

Figure 2

Considerations for NLP in Education.



1.9. Practical Considerations in the Educational Use of NLP

Beyond ethical considerations, there are other practical challenges that teachers and institutions must address when employing NLP-based tools in education. These concerns include cost, accuracy, and integration with established educational procedures (Kumar & Howard).

1.9.1. Accuracy and Limitations of NLP

NLP systems have become more sophisticated, yet they are not perfect. These systems may nevertheless struggle with complicated linguistic tasks including comprehending idiomatic expressions, context and creative writing. Misinterpretations or mistakes in automated assessments might result in inaccurate feedback or unjust evaluations, negatively impacting students' learning experiences (Liu & Yang, 2020, as cited in Kumar & Howard, 2024). Instructors must be aware of these limitations and employ NLP technologies with traditional assessment methods to deliver a balanced evaluation.

1.9.2. Over-Reliance on Automation

One practical issue is the potential over-reliance on automated tools for evaluation and feedback. While NLP systems can be extremely useful in managing vast amounts of student work, there is a risk that instructors will rely too heavily on these systems, ignoring the nuanced judgment and individualized feedback that only human teachers can provide. Striking a balance between automation and human engagement is crucial for providing students both objective assessments and the personal touch that stimulates deeper learning (Wang & Liu, 2021, as cited in Kumar & Howard, 2024).

1.9.3. Cost and Accessibility

Integrating NLP-based technologies in education can be expensive, especially in institutions with limited resources. Developing and maintaining new technologies into current educational infrastructures may necessitate a large financial investment. Furthermore, accessibility difficulties exist, as not all students may have equal access to the devices or platforms needed to interact with these tools. Ensuring equitable access to technology and resolving the digital divide are critical to prevent increasing existing educational inequities (Smith & Rogers, 2022, as cited in Kumar & Howard, 2024).

1.9.4. Integration into Existing Educational Practices

NLP technologies must be viewed as an addition to, rather than a substitute for, established teaching and assessment approaches. To successfully integrate new tools into the classroom, significant attention must be given to how they can assist existing pedagogical practices. Educators may need to be trained to use NLP tools effectively, and the tools themselves must be adaptable to a wide range of teaching styles and curriculum (Hwang & Lai, 2017, as cited in Kumar & Howard, 2024). Furthermore, educators must be involved in the development and improvement of NLP tools to ensure that they suit the practical needs of both teachers and students.

10. Future Directions for NLP in Education

The future of Natural Language Processing (NLP) in education has enormous potential for innovation and transformation. One significant direction is to improve the accuracy and fairness of NLP tools. Because present systems frequently suffer with prejudice and contextual awareness, future research should concentrate on developing more advanced models that better understand nuances in language, linguistic differences and cultural diversity. This will be required more varied datasets, as well as developments in machine learning approaches to ensure fair and equitable assessments for all students, particularly those from non-dominant language backgrounds (Kumar & Howard, 2024).

Conclusion

Advancements in technology have had a tremendous impact on many parts of society, including language learning. The growing integration of Artificial Intelligence (AI) into various industries has created new opportunities for language teaching. In particular, AI has arisen in the field of English teaching, providing unique solutions for language learning. AI-powered language learning platforms are becoming increasingly popular, offering personalized and engaging learning experiences for learners. These platforms use AI technologies like Natural Language Processing (NLP), to analyze and understand text and speech, allowing developers to get insights into learners' language creation. Furthermore, AI has been used into language assessment and evaluation, providing automated language evaluation systems that can compete favorably with traditional approaches. These AI-powered assessment tools enable rapid and objective evaluations of language proficiency. However, the incorporation of AI in English instruction raises ethical concerns that must be addressed.

Chapter Two: Research Methodology

Introduction

The present chapter is devoted to the practical framework of the study. This investigation adopts a descriptive-correlational approach to evaluate how lexical richness changes in relation to the academic level of EFL students. Thus, a corpus analysis was performed on undergraduate students at the English department of 08 May 1945 University, Guelma. The selection of the method was based on the study objectives, which aimed at providing a detailed understanding on how lexical richness, including lexical variety, lexical density and lexical diversity, evolves across academic stages and how it relates to academic development.

2.1. Research Design and Approach

Examining and evaluating lexical richness in higher education necessitates the use of a descriptive-correlational research design along with a quantitative method, implemented through a student-authored corpus analysis as an attempt to answer the research questions as well as to accept or reject the research hypotheses.

Quantitative research is commonly defined as an approach that is centred on testing objective theories by examining the relationship between measured variables. These variables are quantified through instruments that provide numerical data, which are later analyzed through statistical procedures. This design operates through the deductive testing of theories or hypotheses (Creswell & Creswell, 2018). The researchers begin by selecting a theory and formulating a narrow testable prediction, the so-called hypothesis, grounded on that theory.. The following step is collecting data using instruments such as surveys, questionnaires, and tests that measure attitudes or behaviors. Subsequently, the numerical data are analyzed to derive a final conclusion about the tested hypothesis in order to determine whether it is supported or refuted.

Similarly, Barella et al. (2024) posit that quantitative research seeks to reveal relationships, trends and patterns in data with the intention of generalizing the results to a wider population. Quantitative research further aims to yield valid and reliable data in the interest of explaining, predicting or controlling the phenomena under investigation. Ultimately, the quantitative method offers numerous advantages for researchers. It fosters objectivity, enabling unbiased data collection and analysis, as well as the potential generalizability of results, especially when using a representative samples that reflect broader populations. Also, the numerical nature of such data allows an efficient analysis, especially when using computer software, which eventually ensures a better validity and reliability of the results. Owing to the aforementioned benefits and the suitability of the approach to the research topic, we opted for the use of the current research design.

2.2. Sample and Target Population

A target population of 465 first-year and third-year undergraduate students at the Letters and English Language Department, University of 08 May 1945, Guelma was selected during the academic year 2024-2025. The two educational levels were intentionally chosen because the first-year level represents the initial stage of receiving educational instruction at the university, whereas the latter represents the final stage of undergraduate education, denoting a more advanced stage of lexical acquisition. Additionally, the decision was due to the continued exposure of third-year students to academic discourses besides extended writing practices and assignments, such as written and oral presentations, essays and overall research projects. This selection also establishes a comparative basis for the study and allows for observing the developmental changes in lexical richness across the undergraduate educational cycle, in terms of its three major components: lexical density, lexical diversity and lexical sophistication more precisely.

The sample was one of convenience that aligned with the objectives of the research and the sampling method can therefore be classified as convenience purposive sampling. Only those classes' whose teachers expressed willingness to participate in the study were included. The selection was maintained on the basis of practicality and purposeful relevance. Out of 465 students, only 60 were chosen as a sample, 30 out of the 253 first-year students and another 30 out of the 212 third-year students. The thing that allowed for a balanced representation of both educational levels. Another reason behind the selection was the efficiency of data collection, analysis and interpretation within this limited sample size.

2.3. Research Instruments

During the data collection process, a writing task was assigned to both groups of first-year and third-year groups. In the data analysis phase of the investigation, the web-based Lexical Complexity Analyzer (LCA) was employed as a tool for measuring lexical richness in the written texts of students. Microsoft Excel was implemented to organize numerical data obtained by LCA. At a later stage, SPSS software (Statistical Package for the Social Sciences) was used to derive the statistics needed for answering the research questions and for testing the hypotheses based on the output obtained through the use of the LCA. Regarding the reliability and validity of these instruments, the writing task was carefully designed to elicit students' authentic and personal language use, presenting a suitable tool to demonstrate their lexical resources. The choice of the topic ensured content validity by promoting the generation of ample data necessary for effective analysis. Then, the obtained writing samples were analyzed through the LCA, a widely used tool characterized by validity in assessing lexical diversity, lexical sophistication and lexical density. The automated analysis of lexical richness that the tool offers minimizes research biases and errors, leading to consistent results. While SPSS was primarily chosen to

compute statistics, its use enabled a fair and systematic comparison of findings, thereby enhancing the overall reliability and validity of the research.

2.4. Data Collection Procedures

The process of data collection began by administrating the writing task to each of the first-year and third-year groups, whereby it was distributed and collected in a controlled classroom environment during regular sessions at separate intervals, and over a period of two to three weeks. After the collection of the hand-written samples of students, the phase of text transcription took place, which was critically important for preparing the raw data needed for analysis. Equally important, ethical considerations were strictly adhered to at this stage of the study. All participants granted informed consent before engaging in the task. As researchers, we also placed a considerable emphasis on ensuring the anonymity and confidentiality of their responses throughout the research process, in addition to respecting their voluntary participation in the study.

2.5. Corpus Analysis

2.5.1. Aims of the Corpus Analysis

In this research, the corpus analysis is primarily intended to explore and compare lexical richness manifested in the written texts produced by first-year and third-year undergraduate students. It also seeks to identify any measurable differences and common patterns that occur in their vocabulary use. Furthermore, it aims at determining the average scores of lexical richness levels reflected in their writings and offering insights into their language proficiency and academic progression.

2.5.2. Corpus Selection Criteria

As previously detailed, the selection of first-year vs. third-year students was intended to ensure a comparative analysis across the two different academic levels, i.e. the initial and final stage of undergraduate study. Over thirty (30) text samples were collected from

each group. Nevertheless, only samples that met specific criteria were selected. Several texts were eliminated because of were too short to be proceeded by the LCA tool and lacked coherence, including multiple grammar errors that made them difficult to read.

When considering the quantity and type of texts collected, the sixty (60) selected samples were descriptive in nature with some narrative elements due to the nature of the task assigned to them, in which the students had to describe a movie they have watched with a summary of its key events. This descriptive task encouraged them to use varied expressions and adjectives resulting in supporting an effective evaluation of lexical sophistication, lexical diversity and lexical density.

2.5.3. Description of the Corpus

The collected corpora comprised 60 hand-written texts by first-year and third-year students in the classroom. In the given task, both groups were asked to describe the best movie they had ever watched, including what made it special for them, and what feelings did it leave in them, as well as providing a general summary of the plot. Each respondent was given a hard copy of the prompt and wrote their responses in a paragraph form on the same paper. These writings were later on rewrote into a digital version. Another critical notion worth highlighting is that each text sample was treated as an individual unit in the corpus.

The participants of the study are Arabic-speaking learners. Most of them learned French as a second language and English as a foreign language (EFL) within educational settings. There is a considerable variation between their English proficiency levels arising from the differences in exposure to the language.

2.5.4. Corpus Analysis Procedures

The corpus analysis was conducted using the Lexical complexity analyzer (LCA), SPSS software, with a slight assistance from Microsoft Excel, as discussed earlier. To address the first research question—“1) How is lexical richness measured in terms of lexical density, diversity, and sophistication?”—this section provides a detailed account of the commonly adopted metrics in the literature, along with those specifically employed in the present study. The LCA calculated 22 different indices of lexical complexity of students’ written productions including:

- Lexical density (LD).
- Lexical sophistication (LS) which contains: Lexical sophistication I (LS1) and Lexical sophistication II (LS2).

This table illustrates the formulas through which the proxy measures of lexical density and sophistication are calculated (Adapted from Gouider, 2023 b; Lu, 2012; Nasserri & Thompson, 2021):

Table 1

Lexical Density and Sophistication Indices

Measure	Code	Quantification	Example
1-Lexical density	-LD	$-N_{lex}/N$	-Number of content words divided by the number of tokens
2-Lexical sophistication			
a-Lexical sophistication-I	-LS1	$-N_{Slex}/N_{lex}$	- Number of sophisticated content words divided by the number of content words.
b-Lexical sophistication-II	-LS2	$-T_s/T$	-Number of sophisticated lexical types divided by the number of lexical types.

- Lexical variation (LV) counted through:
 - NDW (Number of different words (NDW), NDW (first 50 words) (NDWZ-50), NDW (expected random 50) (NDW-ER50) and NDW (expected sequence 50) (NDW-ES50)),
 - TTR (Type/Token ratio (TTR), Mean Segmental TTR (50) (MSTTR-50), Corrected TTR (CTTR), Root TTR (RTTR), Logarithmic TTR (log TTR) in addition to Uber Index (Uber)),
 - Verb diversity (Verb variation-I (VV1), Squared VV1 (SVV1) as well Corrected VV1 (CVV1)).
 - Lexical word diversity (Lexical word variation (LV), Verb variation-II (VV2), Noun variation (NV), Adjective variation (AdjV), Adverb variation (AdvV) and Modifier Variation (ModV)).

The following table displays the quantification method for the previous metrics (Adapted from Gouider, 2023 c; Lu, 2012):

Table 2

Lexical Variation and Diversity Proxy Measures

Proxy Measures of Lexical Variation	Code	Quantification Method	Explanation
-Lexical variation			
1-Number of different words (expected random 50)	-NDW– ER50	-Mean T of 10 random 50-word sample	-The average number of types in 10 random samples of 50 words.
2-Number of different words (expected sequence 50)	-NDW– ES50	-Mean T of 10 random 50-word sequence with random starting points	- The average number of types in 10 random samples of 50 words sequences with random starting points.
-Lexical verb and word diversity			
1-Verb diversity			
a-Verb Variation-I	- VV1	- $T_{\text{verb}}/N_{\text{verb}}$	- N of verb types /N of verbs.

b-Squared VV1	- SVV1	$-T_{verb}^2 / N_{verb}$	- Squared N of verb types/ N of verbs.
c-Corrected VV1	- CVV1	$-T_{verb} / \sqrt{2N_{verb}}$	- N of verb types/ squared root of verbs' N.
d- Verb variation-II	- VV2	$-T_{verb} / N_{lex}$	- N of verb types / N of lexical words.
2-Lexical word diversity			
a-Lexical word variation	- LV	$-T_{lex} / N_{lex}$	- N of lexical types/ N of lexical words.
b- Noun variation	- NV	$-T_{noun} / N_{lex}$	- N of Noun types/ N of lexical words.
c- Adjective variation	- AdjV	$-T_{adj} / N_{lex}$	- N of adjective types/ N of lexical words.
d- Adverb variation	- AdvV	$-T_{adv} / N_{lex}$	- N of adverb types/ N of lexical words.
e- Modifier variation	- ModV	$-(T_{adj}+T_{adv}) / N_{lex}$	- N of Noun types/ N of lexical words.

The data analysis process followed an organized procedure. Firstly, the hand-written samples were transcribed into digital text format. Second, the transcribed texts were input into the LCA computational tool to calculate the aforementioned lexical complexity metrics, providing detailed numerical data about each learner's linguistics profile based on each submitted text. Thirdly, the obtained results from LCA were carefully structured and organized with Excel. Then, they transferred to SPSS to perform statistical analysis of data. The software's principal function involved calculating descriptive statistics such as the mean (average) scores and standard deviations, as well as providing inferential statistics to decide whether the differences observed between the groups were significant and valid or not.

Conclusion

The study employed a systematic methodology using Lexical Complexity Analyzer (LCA) to measure lexical richness across different metrics, supported by Excel for data organization and SPSS for deep statistical analysis. These chosen tools ensured comprehensive as well as reliable data processing and analysis in addressing research questions. The methodology also ensured reliable measurement of lexical diversity, lexical density and lexical sophistication across the two student groups.

Chapter Three: Results & Discussion

Introduction

The following part of the chapter is devoted to the findings obtained from the analysis of lexical richness in the writings of first-year and third-year undergraduate learners through the SPSS software. The results are presented in tables and organized to reflect the main lexical richness indices measured, shedding light on both key patterns, tendencies and statistical outcomes. The results are structured into three main parts: descriptive statistics, inferential statistics and effect size of the differences.

3.1. Descriptive Statistics

In the analysis of descriptive statistics, the primary focus is put on the obtained mean scores and standard deviations (SD). The mean scores serve to answer the second research question “2) What are the average scores of lexical richness embedded in the spontaneous writings of both groups?”. The latter (SD) identifies how much the scores vary from the average ones. Descriptive statistics in this section will be used also to answer the third research question “3) How do the writings of first-year and third-year EFL students differ in terms of vocabulary use?” through the interpretation of the differences between the obtained scores pertaining to the three aspects of lexical richness targeted in this study. The fourth research question “4) How do academic performance relate to lexical richness development in higher education?” will be also addressed through the elaboration on the statistics provided in the tables below.

Table 3*Descriptive Statistics for Lexical Density and Lexical Sophistication Scores*

	Level	Numbe r	Mean (Average score)	Stanard Deviation	Std. Error Mean
LD	1 st year	30.00	0.22	0.01	0.00
	3 rd year	30.00	0.22	0.01	0.00
LS1	1 st year	30.00	0.72	0.29	0.05
	3 rd year	30.00	0.83	0.02	0.00
LS2	1 st year	30.00	0.59	0.22	0.04
	3 rd year	30.00	0.69	0.01	0.00

The table presents the descriptive statistics for lexical density (LD), lexical sophistication 1 (LS1) and lexical sophistication 2 (LS2) across both groups of first-year and third-year. As shown in the table, for lexical density, both groups are scored identically with an average score of 0.22 and an SD of 0.01. This suggests that both groups use a similar proportion of content or lexical words relative to the total number of words in their texts, although the differences in their vocabulary use, the balance between the content words and tokens remains consistent across the academic levels. Concerning lexical sophistication 1, third-year students showed a higher mean score of 0.83 compared to 0.72 of the first-year group, whereas the SD of third-year (0.02) was less than the first-year deviation (0.29). The same trend appeared in lexical sophistication 2, where mean scores of third-year (0.69), with a low SD of 0.01, were higher than the first-year ones (0.59), which showed a large SD of 0.22. This indicates that third-year learners highly use a sophisticated vocabulary in a consistent manner compared to their first-year peers, for instance, they employed the words “cinematography”, “Immersive” and “post-apocalyptic” in their texts,

which are classified as advanced and low-frequent words according to Oxford Learner's Word List.

Table 4

Descriptive Statistics of Lexical Variation Scores for First-year and Third-year Students

Index	Level	N	Mean (Average score)	Standard. Deviation	Std. Error Mean
NDW	1 st year	30.00	25.53	1.04	.19
	3 rd year	30.00	25.77	1.01	.18
NDWZ	1 st year	30.00	18.23	1.25	.23
	3 rd year	30.00	18.23	1.45	.27
NDWERZ	1 st year	30.00	18.40	.48	.09
	3 rd year	30.00	18.43	.74	.14
NDWESZ	1 st year	30.00	18.21	.58	.11
	3 rd year	30.00	18.30	.60	.11
TTR	1 st year	30.00	.05	.01	.00
	3 rd year	30.00	.03	.01	.00
MSTRR	1 st year	30.00	.37	.01	.00
	3 rd year	30.00	.36	.01	.00
CTTR	1 st year	30.00	.78	.11	.02
	3 rd year	30.00	.65	.11	.02
RTRR	1 st year	30.00	1.10	.15	.03
	3 rd year	30.00	.93	.15	.03
LOGTTR	1 st year	30.00	.52	.02	.00
	3 rd year	30.00	.49	.02	.00
UBER	1 st year	30.00	5.64	.08	.02
	3 rd year	30.00	5.66	.07	.01
VV1	1 st year	30.00	.05	.02	.00
	3 rd year	30.00	.04	.01	.00
SVV1	1 st year	30.00	.05	.02	.00
	3 rd year	30.00	.04	.01	.00
CVV1	1 st year	30.00	.16	.03	.00
	3 rd year	30.00	.13	.02	.00

LV	1 st year	30.00	.03	.01	.00
	3 rd year	30.00	.02	.01	.00
VV2	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
NV	1 st year	30.00	.02	.01	.00
	3 rd year	30.00	.01	.01	.00
ADJV	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
ADV	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
MODV	1 st year	30.00	.22	.01	.00
	3 rd year	30.00	.22	.01	.00

The table reveals the lexical diversity measures for both groups of first-year and third-year. The NDW (Number of Different Words) index was 25.53 for the first-year and 25.77 for the third-year group, reflecting results that were closely aligned. However, the Type-Token Ratio (TTR) of first-year (0.05) was slightly higher than third-year (0.03), depicting more diverse vocabulary relative to the text length compared to the third-year level. Indices such as Mean Segmental TTR (MSTTR), Corrected TTR (CTTR), Root TTR (RTTR), Logarithmic TTR (LOGTTR), Verb variation-I (VV1), Squared VV1 (SVV1), Corrected VV1 (CVV1), Lexical word variation (LV) and Noun variation (NV) showed higher but approximate average scores for first-year learners, while the other measures remained similar in both groups. In conclusion, first-year students demonstrated greater lexical variation compared to third-year students.

3.2. Inferential statistics

Table 3 highlights the inferential statistics of the 24 metrics of lexical richness obtained by the software. This data is used to test the hypotheses presented in the introduction of this dissertation. Before discussing the results, it is essential to note that

statistical significance serves to measure the significance of the differences between the lexical richness scores of the two groups. The calculated p-value represents the criterion based on which the discrepancy between scores is said to be significant or not. In other words, it reveals the likelihood with which such a difference can be obtained by chance. The Two-sided P-value that is highlighted in bold in the table below is chosen for reporting the significance of the difference instead of the one-sided P-value because of the nature of the hypotheses. If the p-value was below 0.05 (<0.05), the observed differences are considered statistically significant and unlikely to occur by chance. In contrast, a p-value above the alpha value that is 0.05 (>0.05) indicates that the differences may be the result of a random chance and not statistically significant.

Table 5*Inferential Statistics Related to Lexical Richness Scores*

Independent Samples Test											
		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	T	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
LD	Equal variances assumed	1.22	.27	.77	58.00	.22	.45	.00	.00	.00	.01
	Equal variances not assumed			.77	54.99	.22	.45	.00	.00	.00	.01
LS1	Equal variances assumed	20.15	<.001	-1.96	58.00	.03	.06	-.10	.05	-.21	.00
	Equal variances not assumed			-1.96	29.38	.03	.06	-.10	.05	-.21	.00
LS2	Equal variances assumed	22.26	<.001	-2.31	58.00	.01	.02	-.09	.04	-.17	-.01
	Equal variances not assumed			-2.31	29.12	.01	.03	-.09	.04	-.18	-.01
ND W	Equal variances assumed	.05	.83	-.88	58.00	.19	.38	-.23	.26	-.76	.30
	Equal variances not assumed			-.88	57.93	.19	.38	-.23	.26	-.76	.30
ND WZ	Equal variances assumed	.79	.38	.00	58.00	.50	1.00	.00	.35	-.70	.70
	Equal variances not assumed			.00	56.73	.50	1.00	.00	.35	-.70	.70
ND WE RZ	Equal variances assumed	3.67	.06	-.19	58.00	.43	.85	-.03	.16	-.35	.29
	Equal variances not assumed			-.19	49.61	.43	.85	-.03	.16	-.35	.29

ND WES Z	Equal variances assumed	.10	.75	-.57	58.00	.29	.57	-.09	.15	-.39	.22
	Equal variances not assumed			-.57	57.94	.29	.57	-.09	.15	-.39	.22
TTR	Equal variances assumed	.25	.62	4.39	58.00	<.001	<.001	.01	.00	.01	.02
	Equal variances not assumed			4.39	57.75	<.001	<.001	.01	.00	.01	.02
MST RR	Equal variances assumed	.52	.48	1.60	58.00	.06	.11	.00	.00	.00	.01
	Equal variances not assumed			1.60	57.52	.06	.11	.00	.00	.00	.01
CTT R	Equal variances assumed	1.09	.30	4.52	58.00	<.001	<.001	.13	.03	.07	.18
	Equal variances not assumed			4.52	58.00	<.001	<.001	.13	.03	.07	.18
RTR R	Equal variances assumed	1.19	.28	4.54	58.00	<.001	<.001	.18	.04	.10	.26
	Equal variances not assumed			4.54	58.00	<.001	<.001	.18	.04	.10	.26
LOG TTR	Equal variances assumed	1.22	.27	4.53	58.00	<.001	<.001	.03	.01	.02	.04
	Equal variances not assumed			4.53	57.99	<.001	<.001	.03	.01	.02	.04
UBE R	Equal variances assumed	.29	.59	-.72	58.00	.24	.47	-.01	.02	-.05	.03
	Equal variances not assumed			-.72	55.63	.24	.47	-.01	.02	-.05	.03
VV1	Equal variances assumed	.78	.38	4.92	58.00	<.001	<.001	.02	.00	.01	.03
	Equal variances not assumed			4.92	55.16	<.001	<.001	.02	.00	.01	.03
SVV 1	Equal variances assumed	.78	.38	4.92	58.00	<.001	<.001	.02	.00	.01	.03
	Equal variances not assumed			4.92	55.16	<.001	<.001	.02	.00	.01	.03
CVV 1	Equal variances assumed	.09	.77	5.11	58.00	<.001	<.001	.03	.01	.02	.04
	Equal variances not assumed			5.11	57.06	<.001	<.001	.03	.01	.02	.04
LV	Equal variances assumed	.03	.87	4.27	58.00	<.001	<.001	.01	.00	.01	.02
	Equal variances not assumed			4.27	58.00	<.001	<.001	.01	.00	.01	.02
VV2	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001	.00	.00	.00	.01
	Equal variances not assumed			3.49	36.90	<.001	.00	.00	.00	.00	.01
NV	Equal variances assumed	.05	.83	3.00	58.00	.00	.00	.00	.00	.00	.01
	Equal variances not assumed			3.00	57.57	.00	.00	.00	.00	.00	.01
ADJ V	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001	.00	.00	.00	.01
	Equal variances not assumed			3.49	36.90	<.001	.00	.00	.00	.00	.01
AD VV	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001	.00	.00	.00	.01
	Equal variances not assumed			3.49	36.90	<.001	.00	.00	.00	.00	.01
MO DV	Equal variances assumed	.49	.48	5.21	58.00	<.001	<.001	.01	.00	.00	.01
	Equal variances not assumed			5.21	56.49	<.001	<.001	.01	.00	.00	.01

The p-value of the lexical density (LD) index was 0.45, indicating that the difference in the score between first-year and third-year levels was likely obtained by

chance. This leads to the rejection of the alternative hypothesis and the adoption of the null hypothesis instead. That is to say, there is no significant difference in the mean scores of lexical density reflected in the written output of first-year and third-year EFL students. LS1, with a p-value of 0.06, was above the typical 0.05 value, this suggests that the variation was not statistically significant, but very close to the threshold of statistical relevance. The significance of the difference between the two groups when it comes to lexical sophistication was more evidenced through the use of LS2 as an index, where a p-value of 0.02 was reported. The latter result signifies a meaningful increase in lexical sophistication among third-year students and therefore provides sufficient evidence to reject the null hypothesis in favor of the alternative. In other words, there is a significant difference in the mean scores of lexical sophistication reflected in the compositions of the two groups. The findings fall in line with the common belief that learners' vocabulary will be more sophisticated as they progress in their academic journey.

In opposition, several Indices of lexical variety(e.g. TTR, CTTR, RTTR), as well as some verb and word variation indices, revealed highly significant discrepancies (p-value of 0.001) where first-year students proved to be superior in diversifying the vocabulary used while writing. Given the p-value is way below the significance level, the null hypothesis is rejected and the alternative one is accepted. That is, there is a significant difference in the mean scores of lexical diversity and variation in the performance of the two groups.

The overall results showed that lexical density stays almost the same as students move forward in their academic career, sophistication tends to be higher with students who belong to an upper academic level while lexical diversity is greater among the first-year students who have lesser exposure to EFL.

3.3. Effect Size of the Differences

The major function of the effect size is that it shows the practical significance of the differences in lexical richness between first-year and third-year learners through the “Cohen’s d” value, in other words, it indicates how meaningful the differences between the lexical richness scores of the two groups. If Cohen’s d value is around 0.2, it demonstrates a small effect that is not noticeable in the real world. A d value of around 0.5 represents a medium effect that is moderately noticeable. While a values of 0.8 denotes a large effect that is obvious and important, meaning a strong difference that has a meaningful impact. A negative Cohen’s d means the third-year group has a higher average than the first-year group.

Table 2
Effect Sizes of the Difference between the two Groups

		Independent Samples Effect Sizes			
		Standardizer ^a	Point Estimate (d)	95% Confidence Interval	
				Lower	Upper
LD	Cohen's d	.01	.20	-.31	.70
LS1	Cohen's d	.21	-.50	-1.02	.01
LS2	Cohen's d	.16	-.60	-1.11	-.08
NDW	Cohen's d	1.02	-.23	-.73	.28
NDWZ	Cohen's d	1.36	.00	-.51	.51
NDWERZ	Cohen's d	.62	-.05	-.55	.46
NDWESZ	Cohen's d	.59	-.15	-.65	.36
TTR	Cohen's d	.01	1.13	.58	1.68
MSTRR	Cohen's d	.01	.41	-.10	.92
CTTR	Cohen's d	.11	1.17	.61	1.71
RTRR	Cohen's d	.15	1.17	.62	1.72
LOGTTR	Cohen's d	.02	1.17	.62	1.72
UBER	Cohen's d	.08	-.19	-.69	.32
VV1	Cohen's d	.01	1.27	.71	1.82
SVV1	Cohen's d	.01	1.27	.71	1.82
CVV1	Cohen's d	.02	1.32	.76	1.88
LV	Cohen's d	.01	1.10	.56	1.64
VV2	Cohen's d	.00	.90	.37	1.43
NV	Cohen's d	.01	.77	.25	1.30
ADJV	Cohen's d	.00	.90	.37	1.43
ADV	Cohen's d	.00	.90	.37	1.43
MODV	Cohen's d	.00	1.35	.78	1.90

Lexical density (LD) showed a small effect size (Cohen's $d=0.2$). Both lexical sophistication measures, LS1 and LS2, revealed a medium negative effect ($d= -0.5, -0.6$). This indicates that third-year students have moderately higher lexical sophistication than first-year students. Moreover, different lexical variation indices, such as TTR (0.13), CTTR (1.17), RTTR (1.17), VV1 (1.27) and LV (1.10), presented a large effect, therefore signifying greater lexical variety among first-year students.

3.4. Summary and Discussion of Corpus Analysis

The descriptive statistics demonstrated that lexical density, calculated as the proportion of lexical items (nouns, verbs, adjectives) to the total number of words (tokens), is remarkably stable across both first-year and third-year groups. Each level showed an average score of 0.22, with a low standard deviation (0.01), the latter shows that the obtained scores were characterized by low variability. However, these outcomes were not statistically significant ($p=0.45$) and corresponded to a negligible effect size ($d=0.20$), denoting that the difference between the density of the texts of the two groups is unlikely to be distinguishable by an average reader. The findings confirm that the balance between lexical words (or items) and the overall number of words remains consistent across the academic levels.

Regarding lexical sophistication, third-year students exhibited higher lexical sophistication (Av. LS1= 0.83, Av.LS2= 0.59) compared to their first-year peers (Av.LS1=0.72, Av.LS2= 0.59). The differences approached the threshold statistical significance ($p=0.02$ for LS1 and $p=0.06$ for L.S2) and they are supported by moderate to large negative size effect (Cohen's $d=-0.5$, Cohen's $d=-0.6$) denoting a meaningful progress in lexical sophistication which increases with academic experience. Here, it is worth noting that The negative sign (Cohen's $d=-0.5$, Cohen's $d=-0.6$) reflects the group coding in SPSS rather than a directional interpretation of the results. These outcomes are

consistent with the findings of Ha (2019), who conducted a study on a group of EFL Korean undergraduate learners, which concluded that lexical sophistication is the most influential factor determining writing quality and it is strongly correlated with greater language proficiency and, therefore greater lexical richness.

In contrast to lexical sophistication, lexical diversity metrics, particularly Type/Token ratio (TTR), Corrected TTR (CTTR), Root TTR (RTTR) and others, underscored that the first-year group manifests higher lexical variety than the third-year group. Moreover, the majority of differences obtained were statistically significant with a p-value of 0.01 which is below the 0.05 cutoff. Also, most of the lexical variation measures showed a large size effect (Cohen's d value >0.8), illustrating that first-year undergraduate learners tend to employ a wider range of vocabulary types, which reflects more diverse and richer language at the entry-level. The expected reason behind such outcomes is linked to the sensitivity of TTR measure to text length. In a longer text sample, the number of repeated words rises; hence, exhibiting a lower TTR value because the proportion of unique words (types) to the total words (tokens) decreases. This limitation is evidenced in the written texts produced by third-year students, which were mostly longer in length, in comparison to first-year students. This directly affirms the research findings of Johansson (2008).

To sum up, the findings of the study revealed that lexical density remained essentially stable across the academic levels. This consistency demonstrates that both groups of first-year and third-year used a comparable proportion of lexical items, regardless of academic progression. Therefore, the null hypothesis is supported. Whereas, third-year students showed greater lexical sophistication that was reflected in their use of more low-frequent vocabulary in their written productions. The findings indicate that academic experience and improved language proficiency contribute to increased

vocabulary sophistication. That is to say, the alternative hypothesis is accepted and the null hypothesis is refused. Contrastingly, first-year students exhibited greater lexical diversity than their third-year counterparts, as evidenced earlier, which confirms the alternative hypothesis and denies the null hypothesis. In conclusion, the results partially support the research hypothesis. There is a significant difference in lexical richness only in terms of lexical sophistication.

Conclusion

As presented in the chapter, the findings indicated a significant difference in lexical sophistication (LS2). as well as in lexical diversity, while lexical density did not show any significant difference. These results were subsequently analyzed and interpreted in detail, providing a deeper exploration of how academic progression can affect lexical richness and its relation to language proficiency.

General Conclusion

In summation, the findings of the current study indicated that lexical density remained consistent across first-year and third-year levels, denoting that academic progression isn't necessarily related to the incorporation of a higher volume of content words. From another angle, third-year students showed a higher level of sophistication, characterised by the use of low-frequency vocabulary. This reflects the role of academic experience and improved language proficiency in more sophisticated lexical choices. Conversely, first-year students presented a higher degree of lexical variety, making use of a broader unique vocabulary range than their third-year counterparts.

Pedagogical Implications

In light of the study outcomes, it is clear that lexical richness plays a paramount role in language proficiency. Building on this, teachers have to employ teaching strategies that enhance lexical diversity, lexical density, as well lexical sophistication. These indices are quite essential for a student's language development. Hence, to promote this progress, educators can implement different vocabulary exercises that are designed specifically to encourage learners to use more varied and advanced vocabulary.

Equally important, since first-year students exhibited higher lexical diversity compared to their third-year peers, teachers should incorporate more activities that encourage the students to use a broader range of vocabulary. Such activities may include writing prompts that encourage students to use diverse expressions, as well as focusing on reading tasks aimed at lexical variation.

Regarding curriculum and classroom practices, educators and instructors need to pay attention to the integration of more opportunities for students to engage with academic texts that promote lexical richness. Furthermore, regular assessments of student's lexical

baggage, tailored with instructional activities, can help in ameliorating lexical sophistication throughout their academic journey.

Limitations of the Research

This study encountered several drawbacks that slowed down the progress of the work. Some were addressed, while others were not.

- This study compares different levels (first and third years) at a single point in time rather than tracking the same students longitudinally. Therefore, it cannot fully account for individual developmental trajectories or external factors that influence lexical growth over time.
- We faced challenges in collecting the data needed for the research as some teachers refused to collaborate, due to time constraints, claiming that timing was unsuitable for their schedules.
- While there are many resources examining lexical richness indices, there was a limited body of work devoted specifically to the concept of lexical richness as a whole.
- Although LCA provided useful quantitative measures, however, it may not fully capture the qualitative aspects of lexical richness.
- The timing of the second-year master students' training, which was conducted in high/middle schools, was inconvenient and slowed down the process. It would have been better if it had been planned during the third-year program.
- Time constraints: Searching and selecting the appropriate sources took a long time.
- Accessing many essential sources for this investigation required financial resources that were beyond the means of the present researchers.

- Reading the students' texts was challenging because they were handwritten and also contained a lot of spelling errors which required correction. Therefore, some of these text samples were excluded from the corpus.
- The study relied on written items that may not fully reflect students' overall lexical proficiency such as spoken vocabulary, receptive knowledge, and pragmatic usage.

Suggestions for Future Research

In addition to the topic of this study, which investigated the evaluation of lexical richness development in higher education through the use of artificial intelligence, various additional issues in this area should be addressed to support the findings and add missing points. Below are some preferred suggestions that may assist future researchers.

- Longitudinal studies that follow the same group of students over multiple academic years would provide deeper insights into the trajectories of lexical development. These designs would allow researchers to monitor individual growth patterns and better understand the factors that contribute to lexical development over time.
- Using mixed-methods research designs that combine quantitative and qualitative data can enrich the analysis and interpretation of lexical development. For example, interviews, classroom observations, and teacher feedback can provide valuable contextual information to complement these statistical findings.
- Further studies should try to account for individual differences among learners, such as motivation, prior exposure to the language, learning strategies, and language use in extracurricular activities. Including these variables would allow for a more accurate examination of the factors influencing the development of lexical richness.
- While automated tools such as the Language Complexity Analyzer (LCA) are valuable for quantitative analysis, future researchers may benefit from combining

automated measures with qualitative analysis.

- Future studies may explore how AI-based tools, such as intelligent tutoring systems, adaptive learning platforms, and advanced vocabulary analysis software, can support and measure lexical development.

Through this research, an effort was made to cover all areas related to the essential elements of lexical richness and AI in language evaluation. The primary goal of this research was to clarify any barriers to understanding and contribute to advancing knowledge, with the hope that this modest research will provide a solid foundation for future research in this area.

References

- Analyze My Writing. (n.d.). Lexical density. Analyze My Writing. Retrived from <https://www.analyzemywriting.com/>
- Ayadi, M. (2023). Lexical richness and syntactic complexity as predictors of academic writing performance. *Journal of English Studies in Arabia Felix*, 2(1), 23-33. <https://doi.org/10.56540/jesaf.v2i1.43>
- Barella, Y., Mustami, A. F., Mustami, M. K., Rahman, U., & Alajali, H. M. A. (2024). Quantitative methods in scientific research. *Jurnal Pendidikan Sosiologi dan Humaniora*, 15(1), 281-287. [10.26418/j-psh.v15i1.71528](https://doi.org/10.26418/j-psh.v15i1.71528)
- Bestgen, Y. (2024). Measuring lexical diversity in texts: The twofold length problem. *Language Learning*, 74(3), 638–671. <https://doi.org/10.1111/lang.12630>
- Chipere, N., Duran, P., Malvern, D., & Richards, B. J. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan. <https://doi.org/10.1057/9780230511804>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications. https://spada.uns.ac.id/pluginfile.php/510378/mod_resource/content/1/creswell.pdf
- Elgobshawi, A. E., & Aldawsari, M. (2022). Lexical density as an improvement indicator in the written performance of EFL majors. *International Journal of English Language and Literature Studies*, 11(4), 181–190. <https://doi.org/10.55493/5019.v11i4.4668>

- Erando, M. R. A., & Fortes, A. C. G. (2024). Lexical richness and features in the journal entries of the L2 learners. *International Journal for Multidisciplinary Research (IJFMR)*, 6(3), 1-13. <https://doi.org/10.36948/ijfmr.2024.v06i03.22338>
- Gouider, I. (2023, March 8-9 a). *Assessing the lexical sophistication of Algerian EFL academic writings* [Paper presentation]. The First National Hybrid Conference on New Teaching Trends in Education, ICT and Economics. Illizi University Centre, Algeria. https://www.researchgate.net/publication/372559299_Assessing_the_Lexical_Sophistication_of_Algerian_EFL_Academic_Writings
- Gouider, I. (2023 b). *Exploring the Lexical Diversity of Texts Through the Use of Modern Technologies: An Automated Corpus Analysis* [Paper presentation]. Modern Educational Technology for Quality and Transformative Learning: Local Needs and Global Challenges. Batna University, Algeria.
- Gouider, I. (2023 c). *The automatic measurement of the lexical complexity of Algerian EFL academic writings: A comparative corpus-based study* [Paper presentation]. In National Conference on the Uses of Artificial Intelligence as a Guarantee for the Quality of Higher Education and Scientific Research, Algiers University, Algeria.
- Ha, H. S. (2019). Lexical richness in EFL undergraduate students' academic writing. *English Teaching*, 74(3), 3-28. <https://doi.org/10.15858/engtea.74.3.201909.3>
- Halliday, M. A. K. (1985). *An introduction to functional grammar* (1st ed.). Edward Arnold.
- Han, I., Ji, H., & Ko, Y. (2022). A systematic review of conversational AI in language education: focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48-63. [10.1080/15391523.2022.2142873](https://doi.org/10.1080/15391523.2022.2142873)

- Hider, C. (2024). *Prospecting the Future of Translator Amid AI and Machine Practices and Outputs* [Master's Thesis, Kasdi Merbah Ouargla University]. DSpace at Kasdi Merbah University Ouargla.
- Howard, E., & Kumar, M. (2024). Natural Language Processing in Education: Automating Assessment and Feedback for Language Learners. *Journal of Informatics Education and Research*, 4(3), 1521- 1529.
- Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*. 19(1), 57-84.
- https://www.researchgate.net/publication/278717839_Short_texts_best-fitting_curves_and_new_measures_of_lexical_diversity
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund University, Dept. of Linguistics and Phonetics, Working Papers*, 53, 61–79.
- Kalantari, R., & Gholami, J. (2017). Lexical complexity development from dynamic systems theory perspective: Lexical density, diversity, and sophistication. *International Journal of Instruction*, 10(4), 1-18.
- [10.12973/iji.2017.1041a](https://doi.org/10.12973/iji.2017.1041a)
- Kilic, M. (2019). Vocabulary knowledge as a predictor of performance in writing and speaking: A case of Turkish EFL learners. *PASAA*, 57, 134-164. <http://dx.doi.org/10.58837/CHULA.PASAA.57.1.6>
- Kyle, K. (2019). Measuring lexical richness. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., pp. 454-476). Routledge.
- Kyle, K., & Crossley, S. A. (2014). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 48(3), 1–30. <https://doi.org/10.1002/tesq.194>

- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190–208. [10.2307/41684069](https://doi.org/10.2307/41684069)
- McCarthy, P. M., & Jarvis, S. (2007). Vocab: A theoretical and empirical evaluation. *Language Testing*, 24(4), 459–488. <https://doi.org/10.1177/0265532207080767>
- Nasseri, M., & Thompson, P. (2021). Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing*, 47, 100511. [10.1016/j.asw.2020.100511](https://doi.org/10.1016/j.asw.2020.100511)
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Özdere, M. (2023). The Integration of Artificial Intelligence in English Education: Opportunities and Challenges. *Language Education & Technology (LET Journal)*, 3(2), 137-172.
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press. <https://archive.org/details/assessingvocabul0000read>
- Smith, S. (n.d.). General service list. https://zakiyama-english.com/wp/wp-content/uploads/openends/GSL_frequency.pdf
- Spring, R., & Johnson, M. (2022). The possibility of improving automated calculation of measures of lexical richness for EFL writing: A comparison of the LCA, NLTK, and SpaCy tools. *Language Testing in Asia*, 11(4). <https://doi.org/10.1016/j.system.2022.102770>
- Ure, J. (1971). Lexical density and register differentiation. In G. Perren & J. M. Trim (Eds.), *Applications of Linguistics* (pp. 443-452). Cambridge University Press.
- Yule, G. (1944). *The statistical study of literary vocabulary*. Cambridge University Press. https://pure.mpg.de/rest/items/item_2407784_3/component/file_2622080/content

- Zhang, Y., & Wu, W. (2021). How effective are lexical richness measures for differentiations of vocabulary proficiency? A comprehensive examination with clustering analysis. *Language Testing in Asia*, 11(15), 1-19. <https://doi.org/10.1186/s40468-021-00133-6>

Appendix A: Writing Task

University of 8 May 1945, Guelma

Duration : 20 min

English Department

Task

The question: In a form of a paragraph, talk about the best movie you have ever watched

(what is this movie, what makes it special for you, what emotions did it leave you with ...)

The answer

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Thank you ^-^

Appendix B: Samples of First-Year and Third-Year Students

Samples of First-Year Students

1) I will talk about the movie of The Hills Have Eyes, is a famous horror movie, it tells the story of a family attacked by savage people in a remote area. And makes it special for movie is the fight between good and evil and how people react in tough situation. After watched, I felt tense and scared, but also it makes me think about survival and courage.

2) For me, the best movie I have watched recently is a Korean movie talked about two girls living in the same house, but in different year, and they can only connecting each other in phone so, it started with one of the girls, her dad died and she moved to the house where they used to live. Someday, the other girl start calling her asking for some help because her mother trying to kill her, at first, she did not believe her, but with time she started believing her. They became friends sharing information of the different world and started to change the future with the help of the girl that lives in the past.

3) Harry Potter is the one that is make me I feel the most connected, safe, and nostalgic. It gives me a sense of comfort and magic like escaping to a world understand me. It lets me live the moment with a special kind of magic, as if I am a part of the story.

4) My favorite movie was about alone man in the desert, it is my favorite because I'm sick of western movies and their propaganda they are trying to push into my brain. The man had a weapon, with that weapon he collected friends and not enemies because he never used that weapon nor he applied its possession, when I was watching that movie at first it wasn't that interesting so I felt the burden around me. But in the last half hour I realized it all, my mind now thinks differently or better than ever.

5) My favorite movie is called Thriller because I like scary movies about serial killers and crimes, it's a special one for me because the main character in this movie cut heads and feet, hands, etc. I feel very happy and comfortable watching it because I imagine myself the same killer and doing the same things with evil people in my real life, I really like to watch suffering and bleeding it a real fun.

6) My favorite movie is The Notebook. It's a famous movie. The main characteristic are Ryan Gosling, Rachel McAdams, James Marsden, and Joan Allen, is a romantic story and love story between Noah and Allie, enduring through time and challenges. I feel mix of love, sadness, and inspiration when I watched this movie and it's my favorite one.

7) I have watched so many movies, but among the best movies I have ever watched is Mulan, which I watched many times when I was younger. The thing that makes this film special for me is the sacrifice of Mulan for her father when he went to war, and then the love story between Mulan and the leader. Also, how she saved the king from the villain. She was brave, and nothing could stop her from protecting her family.

8) One of the most memorable cinematic experiences I was watching was The Shawshank Redemption. What makes this film special to me is its narrative of hope, perseverance, and friendship. The movie tells the story of two men who meet in prison, become friends, and plan their escape. The film evokes a range of emotions in me from hope to despair and from joy to sadness. This film has had a profound impact on me.

9) For me, the best movie of all time is The Innocent, is a 1993 American drama. The film tells the story of a man who is accused of a crime he did not commit and tries to prove his innocence. The film is distinguished by its beautiful cinematography and outstanding performance by the stars.

10) Squid Game is a South Korean survival thriller series. The show centres around a deadly competition where 456 financially players participate to engage in a series of children's games. There is a doll in the series named Young-Hee. It's giant and sing a horror song. This movie special for me because I like Korean technical works and culture.

11) My favorite movie is Harry Potter because it's kind of imagination movie and that what I like, non-reality movie. And it's very special to my heart because I watched it when I was young with my family, which I made memories with them while watching it and also my favorite sentence is Expelliarmus and when he wants to kill, he says Crucio.

12) My favorite movie since I was a kid is Harry Potter. I really like its vibes because I prefer fantasy films. I really loved Hogwarts school and always wanted to study there and learn magic spells, and be friend with Harry, Ron, and Hermione. I also wish to belong to the house of Gryffindor. Harry Potter make me warm especially in the fall.

13) My best movie is The Beekeeper. It's talking about a man who lives in a small village and takes care with the bees and his old neighbor. One day this old lady died because a corruption network had cheated on her and they took from her all the bank balance. After that he took the revenge from the lady and he had destroyed their whole company. I like this movie because it motivates me to be a good man and trying to protect my people.

14) I am a movie lover, so when it comes to choosing one to talk about is kind a hard, but I make up my mind to speak about The Witch is a Korean movie, not for everyone, because of its creepy story. It talks about an organization that kidnaps kids and apply a crazy barbaric experiment on them, in order to give them kind of supernatural powers. It is runed by a brain surgeon doctor "woman", funded by a mysterious organization, but at a certain point, the funds and money stops, and orders come to her to stop these experiments called "old generation kids", so she commits a very brutal action, which is to kill all the kids,

hundreds of kids, and just burn the place. While all of this is happening, one girl escapes. So, the doctor sends some boy to look for her. But, as escaped, she runs away and she ends up at a farm of two lovely old couple. Life passed and she is eighteen now. Sometimes she suffers from severe headache and more bleed. Events keep going and she participates in a talent show. She sings well and followed it by a magic move. Because of the magic staff, the organization found her and start following her around till they threatened her by her family, so she goes with them. And the doctor meet her again and gave her the cure for her head to win her over. But what they didn't know is she actually the one who found them by doing the trick on the TV. And when she was young, those brutal experiments actually give her powers, so she is the one who found the family and all of it was her doing. And she kills everyone at the end.

15) My best movie is Shazam. The story of the movie is a superhero who fights the power of evil in a different world. and he faces a lot of obstacles and sacrifices. I like this movie because it mix fantasy and reality and its story is very interesting when watching. It make you want to watch another part.

16) One of the best movies I have ever watched is Rapunzel. It's a beautiful and magical story about a girl with long, golden hair who lives in a tower. She doesn't know anything about the outside world until she meets a man who helps her escape. I liked the story because it was full of adventure, love, and fun. The music was also amazing. I watched it many times when I was a child, and it always makes me feel happy.

17) The best series I have ever watched is The Vampire Diaries. It's full of action, romance, and mystery. I love the characters, especially Damon and Elena. The story is about vampires, werewolves, and other supernatural creatures living in a small town. The

relationships and suspense kept me watching for hours. I learned many English expressions from it, and I would love to watch it again.

18) I watched Squid Game, and it was one of the scariest series I've ever seen. It's about many people playing games to win money, but if they lose, they die. The players try to win and survive, even if it means hurting others. The series made me feel scared and suspicious because you can't trust anyone. It taught me that money can make people do terrible things.

19) One of the most interesting stories I have read is Animal Farm. It's a novel, not a movie, but I still want to talk about it. It shows how animals take control of a farm, but then they end up having leaders who are just as bad as the humans. It's about power and control. Even though it's a story about animals, it really talks about politics and human behavior. It made me think a lot.

20) One of my favorite animated movies is Moana. It's about a girl who is brave and loves the ocean. She goes on an adventure to save her island and meets Maui, the strong man, who helps her. I liked the songs, the colors, and the story. Moana is not like other princesses—she is strong and independent. This movie inspired me a lot.

21) The movie Bridge to Terabithia was very emotional for me. It's about two children who imagine a magical world together in the forest. They become best friends and have fun until something very sad happens. This movie made me cry, but it also made me value friendship and imagination. It shows how powerful our dreams and friendships can be.

22) One of the movies that confused me at first was Ekvillain. It's about a man who used to be a criminal but then fell in love with a woman who helped him change. Unfortunately, something tragic happens to her, and he wants revenge. The movie has action, sadness, and deep emotions. I liked it because it showed how love can change someone.

23) I liked *The Great Gatsby* because it was full of emotions and had a deep story. The main character is mysterious and rich, and he throws big parties to impress a woman he loves. The movie talks about love, money, and how people sometimes wear masks. The music and scenes were beautiful. It made me feel sad, but I really liked it.

24) Claude's family movie was one of the best movies I've ever watched. It was funny and showed how family members can be different but still love and support each other. I laughed a lot while watching it, and I also learned how important it is to accept each other's differences. It's a movie I would watch again.

25) The movie *Baby's Day Out* is very funny. It's about a baby who escapes from three men who try to kidnap him. The baby goes on an adventure around the city, and the men keep failing to catch him. I laughed a lot because the baby was so clever and the men were so silly. It's a great movie to watch with family.

26) *Train to Busan* is a Korean movie about zombies. It's very exciting and full of action. A man and his daughter are on a train when suddenly, a zombie virus spreads. They try to survive and help others on the train. I liked the relationship between the father and daughter, and the ending was very emotional. It showed how people can be brave even when they are scared.

27) The best series I have ever watched is *Harry Potter*. It is special for me because it is about imagination and fiction, which makes me escape from reality and live in a world full of magic, suspense, drama, and mystery. I also loved the friendship between the characters. When I remember this series, I remember my childhood and the beautiful moments I lived with it.

28) Actually, I watched a lot of films and series since I was a kid, which helped me to develop my English level. This is just a small note about the benefits of watching films.

So, the best movie I have ever watched is *It Ends With Us*. It is a little bit complicated, but I will give you a small summary. It talks about a woman and her partner who is violent like her father. So, we can call it a toxic relationship, and she can't end it until her friend from childhood appears and tells her that she should get over this relationship.

29) In fact, I watch series instead of movies, but I can tell you about this. My best series are *The Last Kingdom* and *The Vikings*. I love everything related to history, especially British history and civilizations. These series are full of battles, traditions, and powerful characters, and I enjoy learning from them while watching.

30) So first, I have watched a lot of fun movies, but unfortunately, I don't remember their names. But the best one is *Andrew Parris*. It is a new movie. What makes it special for me is the shark mother that gives birth to a lot of sharks. This is such an enjoyable movie, and that's what makes it fun. It made me feel like I wanted to watch it again because it is such a nice film.

Samples of Third-Year Students

1) I have a lot of favorite movies that I have watched. One of them is *A Quiet Place*, a gripping horror film that explores a post-apocalyptic world where sound-sensitive creatures hunt humans. The story follows a family that must live in silence to survive, using sign language to communicate and navigate their daily lives with extreme caution. The film masterfully builds tension through its minimal dialogue and innovative sound design, creating an immersive experience that highlights the themes of sacrifice, family, and survival. The emotional depth and suspenseful moments make it stand out in the horror genre.

2) The best movie I have ever watched is *Almost Famous*. What makes it special for me are the vibes of the movie, the places, the story, and the characters. It's about music and a

rock and roll band, and that's what I'm interested in. That's exactly why the movie is special to me. The emotion the movie left me with is that life is more beautiful than we think, and it can be more challenging and full of unexpected events. Life is more amazing when we take adventures.

3) It is crucial to watch movies that, in a way, affect your day and have a special impact on your life. I am not the kind of person who watches too much, yet I watched that movie and listened to its song, and I still think about it when I watch other movies. The movie is *La La Land* and its famous song *City of Stars*. It is special because it starts with two people who form a special bond and make music together. Each time they meet, the emotions and feelings in the movie remind me that, despite the hard times in life, they keep pushing to be together. Even when they part ways, they continue to respect each other. Listen to the song—it is so special.

4) One of the best movies I have ever watched was an Indian movie called *Sawaari Teri Kasam*. The plot is about a man who was considered a criminal and had a bad reputation since he was a child. A woman, who was shy and isolated, was killed by her father due to a big misunderstanding. The man tried to protect her and return her to her family, but the father rejected them and never held a funeral for her. As they say, because of that event, the man began to change. He fell in love with her, and she helped him become a good man. He later became a lawyer. In the end, they discovered he was not the one who had killed the woman—it was actually done by a new houseworker.

5) I watched an anime called *Attack on Titan*. It is my favorite anime, and it was special to me because it was the first one I ever watched. It's my childhood anime, and it left me confused, shocked, and, of course, sad because of its ending. The story is about three

children whose city is attacked and destroyed by Titans-giant creatures that bring chaos and fear.

6) I watched an anime, Attack on Titan. It is my favorite anime. It was special for me because it was the first one I had ever watched in my life. It is my childhood anime, and it left me confused and shocked, especially because of its ending. It tells the story of three children whose city is attacked by titans that destroy everything.

7) The best movie I have ever watched is Dead Poets Society. This was one of the most magnificent movies ever. It was about poetry, literature, and the artistic human touch a mix between emotions, psychology, and logic. Dead Poets Society is known for its clarity in showing challenges, yet complexity in finding yourself and fulfilling your wishes. It shows how poetry can affect a student's life through a teacher who chose to connect dreams with reality rather than just following the administration's rules. One of the students wanted to be an actor. After struggling with his strict father, he wanted to do something he enjoyed, especially after the teacher encouraged him, though his father refused. The boy killed himself in the end, and the father blamed the teacher, who was then expelled.

8) Prison Break is one of my favorite shows. It's a series, actually. The story is about a group of prisoners planning to escape from prison. It starts when the older brother, Lincoln, is arrested and accused of a crime. Everyone believes he is guilty. His brother, Michael, tries to save him before his execution. Michael purposely commits a crime so he can be arrested and put in the same prison. He had tattooed his entire back with the prison's blueprint. Eventually, he manages to escape with his brother, and they leave the country forever.

9) The best series I have ever watched is Squid Game. It is a horror series about different games in which, if someone fails or runs out of time, they get killed. Many people join this

game because the winner receives a large amount of money. The series is very suspenseful because every person tries to betray even their friends to win the game. It left me with a strong emotion: never trust everybody.

10) One of the best movies I have ever watched is Inception. The story is super cool because it has different layers of dreams, and you never know what's real and what's not. The visuals are amazing, and the action scenes are intense. It's a movie that makes you think and keeps you guessing. That's why I loved it so much.

11) One of my favorite movies is called Little Miss Sunshine. It tells the story of a family of six going on a journey to let the youngest, a little girl named Olive, compete in a beauty pageant. The family members all deal with their own issues and discover each other's struggles, learning to empathize, accept, and support one another during the road trip. This movie is special to me because it honestly shows the everyday personal issues people face and how support from loved ones helps. It left me feeling content with my problems and grateful I'm not alone in my personal battles.

12) One of the best movies I have ever watched is Cinderella, the live-action version. It is about a young girl who lives with her cruel stepmother and stepsisters who treat her badly. Despite everything, Cinderella remains kind and hopeful. One day, the king organizes a grand ball to find a bride for his son. Cinderella wants to go, but her stepmother stops her. She cries until a fairy appears and gives her a beautiful dark blue dress and glass slippers. She goes to the ball and dances with the prince, but must leave by midnight. As she rushes away, she leaves behind one of her slippers. The prince searches the whole kingdom until he finds her, and they get married in the end.

13) The best series I have ever watched is Harry Potter. It is special for me because it is about imagination and fiction, which allows me to escape from reality and live in a world

full of magic, suspense, drama, and mystery. It also shows deep friendship. When I remember this series, I remember my childhood and the beautiful moments I lived.

14) My favorite movie I have ever watched is Little Women. It is about a young woman who loves writing books to provide for her family. She is very strong, helpful to her sisters, and determined to succeed in a society that believes girls should only look for a rich husband. It's a movie full of emotions: love, sadness, loneliness, longing, hate, success—and what makes it special for me is the happy ending.

15) My favorite movie is Saw. It's about a character named John Kramer, also known as the Jigsaw Killer, who captures his victims and places them in life-threatening situations to test their will to live. The story explores themes of morality, survival, and the consequences of one's actions. The movie left me feeling a mix of suspense, fear, and even moral questioning. It made me reflect on life choices and the value of life itself. My favorite phrase in the movie was: "Make your choice".

16) From the best movies that I had watched, Alpha, which is a movie about a boy who survived after falling from a high place and get attacked by wolves besides his merciful personality, when he treated the wolf that he injured it, this later which changed its savage behavior. What makes the movies special is that they did not speak a lot there is no dialogues, but the meaning is revealed. The way that the boy and the wolf survived passing by different challenges and in hard conditions of cold etc. inspired me of how to be patient and never surrender or give up. The movie also shows how to change the bad behavior by the idea of domestication of wolves and how when they treat the animal well, it become useful rather than dangerous.

17) The best movie that I have ever watched was in my childhood, and I still remember it since now. And I think that the majority of people or children have been watched it. The

movie was Home Alone. This movie was special because it was so funny and comedy. Also, the language was simple and clear for all levels and ages. The movie of Home Alone telling us a story of a boy between eight or nine years old, how was so rowdy and his family forgot him in the house when they are traveling to another country. Then, two persons came to the house to steal it because they think it was empty, but they found the boy there and he make prank and trap for them which make the movie full of comedy.

18) The best movie I have ever watched was Divergent. With its three parts, it is science fiction, action, and romance genre. The storyline took a place in the future, where people have been isolated in a limited place. They were told that they are the last people in earth. They were categorized in different classifications based on their personalities and capacities and they were not allowed to be divergent, which means having more than one perspective. You should stick on very specific set of characteristics otherwise, you die. Examples of these categories there are Braves and Peacemakers. The reason why did I choose this movie is my interest in psychoanalysis as someone who really loves classify people into different types and loves reading about personalities. I found it very interesting that I felt they made it for me. Finally, the movie left a good and positive impression in me and I highly recommend it.

19) My favorite movie is Wicked. It talks about the story of Elphaba, the Green Witch, before she becomes an evil witch that is hated by everyone. The movie began with her joining the Shizschool, and there she meets Glinda, the good witch. The movie shows their relation development from hating each other to be best friends. They travel to the Oz world, and then everything changed. The movie contained a lot of emotions, music that stuck in your head, and a little bit of romance. The cinematography and the cast was on point, make it one of the best movies that I have ever watched.

20) Since I have nothing in my mind at the moment, I will talk about a series I watched recently that is definitely going to be on my favorite shows list. It is high potential. It is a detective crime series, composed of thirteen episodes, and this week is the last episode so sad. It talks about Morgan, a cleaning lady in a police station, who has a high IQ level and photography memory. One night she found a crime file and solved it. The next day all the police station was for her, and since then she become a consulting with the major crime section. Each episode there is a new case to solve and in the meantime she is gathering information about her ex-husband who disappeared fifteen years ago. The reason why I love this series is simply because I have always been interested in crimes and detective cases since a young age, with Detective Conan and Scooby-Doo. And good luck with your dissertation!

21) The Hobbit, it is a movie, talks about a short man lives in a small town and he is a writer. He wrote about everything. One day he wanted to write about something new. One day a wizard man comes to him and asked him to join a group that they want to relieve their kingdom from a dragon. The short man rejected at first, then he accepted. He didn't expect to face monsters and magical places. He really enjoyed the journey and he was satisfied at the end he wrote about this journey. The moral from the movie is you have just to be brave and do what you want.

22) Personally, I had not have an elite best movie, but Pride and Prejudice is one of my favorite movies. It is depicted from Jane Austen's book. It talks about two characters, Mr. Darcy and Miss Elizabeth, and their journey. The plot was more about how Mr. Darcy had this picture of Elizabeth and her family and how they are so different. He was from the higher class, while she was from a lower class, labour. What I liked about this movie is the cinematography and the emotions shown, whether love, greed, disgust, etc. For the emotions, it left me with, for some reason, I hated the mom for how she forced her

daughters to marry. I understand it was hard financially before, still, I didn't really like her. On the contrary, I loved the father, who was supporting and understanding of his daughters and didn't really impose anything on them. It's a good one. Give it a watch.

23) The best television series I have ever watched and still watching is Friends. It is a show that includes six main characters. Ross, Rachel, Joey, Phoebe, Chandler, and Monica. They are all in their twenties, living in New York City. The show starts by them gathering in a coffee shop called "Central Perk". Each of them has a different personality, but in a way that makes them the ideal friends group. Like I was saying, each has a different personality. Ross has a Ph.D. in paleontology, his greatest achievement, and never miss a chance mentioning it. Chandler is the funny, sarcastic one that makes jokes out of everything. Rachel is the spoiled, girly one. Phoebe is the weird one who does not care of what others think of her. Joey is not very smart one, but very sweet and caring. Monica is the mother of the group. She is the strict one, and as they call her, the one who holds the group together. Watching Friends was one of the best decisions I have ever made. It helped me a lot, especially when life was not worth living. It helped me go through it. Also, it played a huge part in learning vocabulary.

24) Kalpatış, or Heartbeat, is a Turkish series that tells a story of a naughty girl who lived with her father and stepmother, then she moved to live with her grandfather since she kept making troubles in school, her grandmother succeeded in convincing her to continue school and to be a doctor. Her life changed again when her grandmother died because of a non-caring doctor that caused her death. She continued studying to be a doctor and to work in the hospital that doctor works at to reveal the truth of what happened in that awful bad day.

25) The series that I watched and still influenced me still now was Bruce Lee. I think that most of people know them. What makes it special to this series that is even the hard condition that supposed to in your life can face it in order to achieve or comes to your goal. The actions of this series was full of suspense. When I watched this series, I was feel like I had energy that I can do everything.

26) The title of the movie is Frozen. The movie is about a group of three friends that went to an ice mountain and they got stuck in a zone hanging there, then tried all the ways to escape, then two people died and one still alive, and come back home. What makes it special is that it contains excitement and suspense. I feel sad.

27) The best movie I have ever watched is a movie called Maleficent. It is an old movie about an evil witch called Maleficent who have cursed a princess named Aurora because of some problems happened between Maleficent's kingdom and another kingdom. Although Maleficent cursed Aurora, but she really loved her by time and she sacrificed herself to save her. And this made the movie very special to me and left a long lasting impression on me and helped me to understand what is real love.

28) The name of the movie is Pink. It is an Indian movie. Its story is about three women, Meenal, Falak, and another one, I forget her name. One night, they were attacked by a group of men, and when they tried to defend themselves, the problem got bigger, leading to a battle in court means the women are put on trial for their defend, not the action of men. It is special for me because the lawyer tried to help them as much as possible. And tension increased when the truth is hidden because of the power of money in each court, but in the end, they won that battle after a lot of problems and lies. It left me happy. I was watching with focus the whole movie, wishing that the truth will win.

29) About a month ago, I watched a Korean movie, entitled *Forgotten*. It is one of the best movies I ever watched. It talks about an eighteen-years old boy who finds himself displaced to another house with his family. But, in fact, the story is about a forty years man who is a serial killer living with investigators who played the role of his family, to see how he did kill those people. It is a very interesting movie. What makes it special for me is the ambiguity and mystery. I felt very curious and sad while watching it. It is a wonderful movie. I recommend it to you to watch it. You will enjoy it.

30) The movie of *Trip of Space*, it is scientific and knowledge, i.e., movie describes the ability of God in universe such as stars and different planets. I enjoy when I watch this movie because I learn it a lot of information about astronomy, especially for "Big Bang", theory to describe start shape of universe and how to develop to arrive for today. In addition, the scientists show for how Earth surround about itself and about Sun and how happened change between light and night. Next, this movie know us about nucleus, involves for protons and neutrons and how electrons surround about Pluton. Finally, I hope all people watch this movie because help them to take a lot of information and knowledge about space and they live a good trip when watch this movie and enjoy it for different events that is very suspense.

Appendix C: Lexical Complexity Analyzer Data Processing Steps

Step One: Click the link of the Lexical Complexity Analyzer tool below and choose LCA from the home page.

<https://aihaiyang.com/software/l2sca/>

Figure 3

Web-Based Lexical Complexity Analyzer

Haiyang Ai
Login | Register

Home LCA L2SCA WeCLECA

Web-based Lexical Complexity Analyzer

The **Lexical Complexity Analyzer (LCA)**, developed by **Professor Xiaofei Lu** at The Pennsylvania State University, is a tool that allows language teachers and researchers to analyze the lexical complexity of written English language samples, using 25 different measures of lexical density, variation and sophistication proposed in the first and second language development literature. The software runs on UNIX-like (LINUX, MAC OS, or UNIX) systems, and require the input texts to be part-of-speech (POS) tagged and lemmatized. This likely calls for familiarity of the command-line interface as well as some programming skills (e.g., part-of-speech tagging and lemmatization). The web-based interface to LCA, available on this website, eliminates the need for the command line interface and streamlines the above-mentioned natural language processing (NLP) processes, and generate the results in just a few clicks away.

Web-based LCA: Single Mode

The single mode allows you to analyze a single text (or compare two texts) for selected lexical complexity measures. You may choose to see the results of any or all of the 25 indices, and the system will create a graphical representation to visualize the results. Additionally, you may enter another text in order to compare their lexical complexity.

Web-based LCA: Batch Mode

The batch mode allows you to analyze lexical complexity of written English samples up to **200 files at a time**. The results will be a CSV file that can be subsequently imported into spreadsheets or statistical packages for further analysis. Note that the batch mode requires you to **register an account** before using it. The registration is free and take less than a minute.

By using the web-based software described above, you are acknowledging that you agree to be legally bound and to abide

Step Two: Choose the Web-based LCA: Single Mode.

Figure 4

LCA Single Mode

Haiyang Ai
Login | Register

Home LCA L2SCA WeCLECA

Web-based Lexical complexity analyzer - Single Mode

The Single Mode of the web-based Lexical Complexity Analyzer takes an English text as input and computes 25 indices of lexical complexity of the text. You may choose to see the results of any or all of the 25 indices, and the system will create a graphical representation to visualize the results. Additionally, you may enter another text in order to compare their lexical complexity. Please note that each text should have a minimum of 50 words and a maximum of 10,000 words. If you have multiple files to be analyzed, please use the **Batch Mode**. By accessing and using the Lexical Complexity Analyzer, you are acknowledging that you agree to be legally bound and to abide by the **LCA Terms of Service**. If you intend to publish a paper that used the web-based interface to the LCA software, please cite:

- Ai, Haiyang and Lu, Xiaofei (2010). *A web-based system for automatic measurement of lexical complexity*. Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10). Amherst, MA. June 8-12.
- Lu, Xiaofei (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Language Journal*, 96(2):190-208.

Step 1: Enter text #1

Step Three: Inter one text each time.

Figure 5

Input Insertion into the LCA

Step 1: Enter text #1

For me, the best movie I have watched recently is a Korean movie talked about two girls living in the same house, but in different year, and they can only connecting each other in phone so, it started with one of the girls, her dad died and she moved to the house where they used to live. Someday, the other girl start calling her asking for some help because her mother trying to kill her, at first, she did not believe her, but with time she started believing her. They became friends sharing information of the different world and started to change the future with the help of the girl that lives in the past.#1

Enter text #2 (optional)

Step 2: Select indice(s)

Step Four: Select indice(s).

Figure 6

Selection of Lexical Richness Indices

Step 2: Select indice(s)

Lexical density (LD)
Lexical Sophistication
Lexical sophistication-I (LS1)
Lexical sophistication-II (LS2)
Verb sophistication-I (VS1)
Verb sophistication-II (VS2)
Corrected VS1 (CVS1)
Lexical Variation
NDW
Number of different words (NDW)
NDW (first 50 words) (NDWZ-50)
NDW (expected random 50) (NDW-ER50)
NDW (expected sequence 50) (NDW-ES50)
TTR
Type/Token ratio (TTR)
Mean Segmental TTR (50) (MSTTR-50)
Corrected TTR (CTTR)
Root TTR (RTTR)
Bilogarithmic TTR (logTTR)
Uber Index (Uber)
Verb diversity
Verb variation-I (VV1)
Squared VV1 (SVV1)
Corrected VV1 (CVV1)
Lexical word diversity
Lexical word variation (LV)
Verb variation-II (VV2)
Noun variation (NV)
Adjective variation (AdjV)
Adverb variation (AdvV)
Modifier variation (ModV)

Tip: Press CTRL or SHIFT to select multiple indice(s)

Step five: Select English Variety and submit.

Figure 7

Selection of English Variety

ig.com/software/lca/single/

NDW (expected sequence 50) (NDW-ES50)

TTR
Type/Token ratio (TTR)
Mean Segmental TTR (50) (MSTTR-50)
Corrected TTR (CTTR)
Root TTR (RTTR)
Bilogarithmic TTR (logTTR)
Uber Index (Uber)

Verb diversity
Verb variation-I (VV1)
Squared VV1 (SVV1)
Corrected VV1 (CVV1)

Lexical word diversity
Lexical word variation (LV)
Verb variation-II (VV2)
Noun variation (NV)
Adjective variation (AdjV)
Adverb variation (AdvV)
Modifier variation (ModV)

Tip: Press CTRL or SHIFT to select multiple indice(s)

Step 3: Select English Variety:

British English

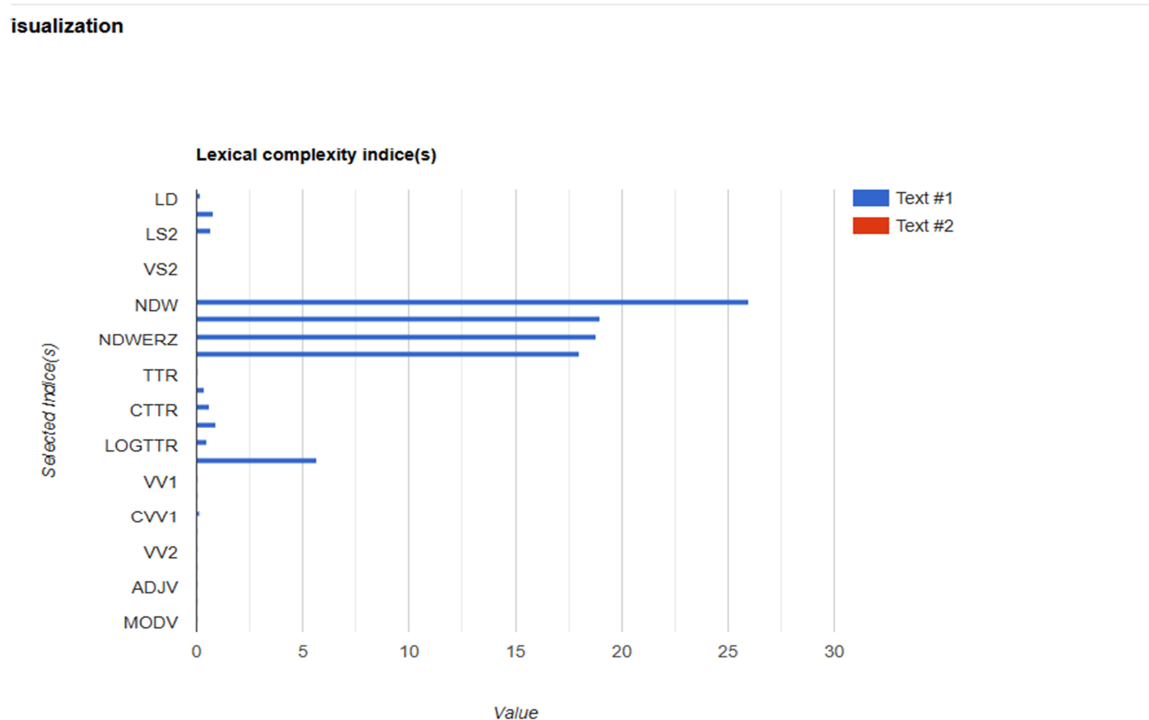
Submit

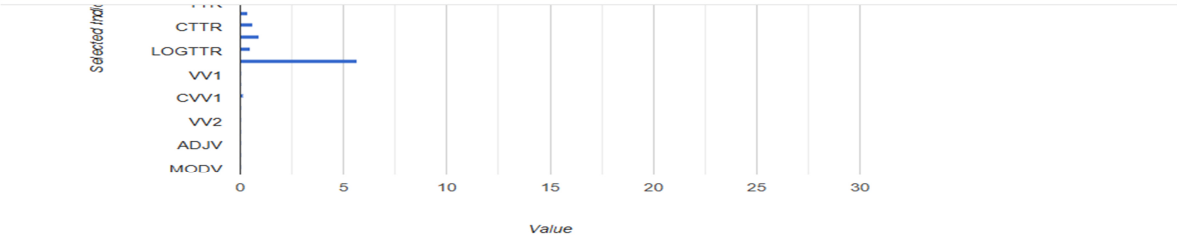
Visualization

Step Six: Results Shown as both Visualization and Numeric Results.

Figure 8

Visualization and Numerical Results





Numeric Results

The following tabular results can be copied and pasted into a plain text file, and subsequently imported into spreadsheet or database software for further processing and analysis.

Text #1

LD, LS1, LS2, VS1, VS2, CVS1, NDW, NDWZ, NDWERZ, NDWESZ, TTR, MSTTR, CTTR, RTTR, LOGTTR, UBER, VV1, SVV1, CVV1, LV, VV2, NV, ADJV, ADVV, MODV

0.21, 0.81, 0.65, 0.00, 0.00, 0.00, 26, 19, 18.80, 18.00, 0.03, 0.36, 0.63, 0.89, 0.48, 5.67, 0.03, 0.03, 0.12, 0.02, 0.01, 0.01, 0.01, 0.01, 0.01

Computed in 0.52450 seconds.

Appendix D: Detailed Statistical Tables

Figure 9

Output from Excel Sheet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	student	LD	LS1	LS2	VS1	VS2	CVS1	NDW	NDWZ	NDWERZ	NDWESZ	TTR	MSTRR	CTTR	RTRR
2	1	0.23	0.83	0.68	0	0	0	25	19	18.6	19.3	0.05	0.38	0.79	1.0
3	2	0.21	0.81	0.68	0	0	0	25	19	18.6	18.1	0.03	0.37	0.61	0.8
4	3	0.21	0	0.04	0	0	0	25	15	18.2	18	0.07	0.36	0.91	1.0
5	4	0.23	0	0.04	0	0	0	25	18	18	17.7	0.04	0.36	0.66	0.6
6	5	0.23	0.87	0.68	0	0	0	25	19	18.3	18.8	0.04	0.38	0.74	1.0
7	6	0.26	0.88	0.69	0	0	0	26	19	18.2	17.5	0.05	0.36	0.84	1.0
8	7	0.22	0.82	0.68	0	0	0	25	19	17.8	18.8	0.04	0.38	0.73	1.0
9	8	0.23	0.85	0.69	0	0	0	26	17	18.7	17.7	0.05	0.36	0.77	1.0
10	9	0.22	0.85	0.63	0	0	0	27	18	18.9	18.8	0.07	0.38	0.95	1.0
11	10	0.24	0.91	0.63	0	0	0	30	21	17.8	18.2	0.06	0.37	0.97	1.0
12	11	0.22	0.83	0.69	0	0	0	26	17	17.9	19.3	0.05	0.39	0.83	1.0
13	12	0.24	0.86	0.68	0	0	0	25	19	17.8	17.9	0.05	0.35	0.8	1.0
14	13	0.21	0.81	0.68	0	0	0	25	19	18.4	17.6	0.04	0.36	0.71	1.0
15	14	0.21	0.84	0.69	0	0	0	26	17	18.8	17.5	0.01	0.37	0.39	0.5
16	15	0.22	0.83	0.69	0	0	0	26	19	18.6	19	0.06	0.37	0.88	1.0
17	16	0.22	0.83	0.68	0	0	0	25	19	19.1	17.7	0.04	0.37	0.72	1.0
18	17	0.24	0.86	0.69	0	0	0	26	18	18.5	18.3	0.05	0.38	0.79	1.0
19	18	0.22	0.79	0.69	0	0	0	26	19	18.3	18.3	0.05	0.36	0.81	1.0
20	19	0.21	0.84	0.68	0	0	0	25	18	18.3	17.4	0.04	0.35	0.74	1.0
21	20	0.23	0.83	0.69	0	0	0	26	18	19.2	17.9	0.06	0.36	0.85	1.0
22	21	0.22	0.85	0.68	0	0	0	25	15	19	19.3	0.05	0.38	0.82	1.0
23	22	0.21	0.8	0.68	0	0	0	25	19	17.9	18.1	0.05	0.37	0.81	1.0
24	23	0.21	0.84	0.68	0	0	0	25	19	19.1	18.4	0.05	0.38	0.82	1.0
25	24	0.2	0.77	0.67	0	0	0	24	18	18.4	18.3	0.05	0.37	0.8	1.0
26	25	0.2	0	0.04	0	0	0	25	18	19.2	18.9	0.06	0.37	0.84	1.0
27	26	0.25	0	0.04	0	0	0	26	19	17.7	18.4	0.05	0.35	0.81	1.0

Figure 10*SPSS Dataset*

File Edit View Data Transform Analyze Graphs Utilities Extensions Windows Help

<

Appendix E: Output from SPSS

Note. 1. Standard Deviation (SD): Interpretation:

What it tells you: How much individual scores differ from the average (mean).

A **small SD** suggests that most scores are close to the mean (less variability).

A **large SD** indicates more spread—scores are more dispersed.

Table 3

Holistic Descriptive Statistics

Descriptive Statistics					
	Level	N	Mean (Average score)	Std. Deviation	Std. Error Mean
LD	1 st year	30.00	.22	.01	.00
	3 rd year	30.00	.22	.01	.00
LS1	1 st year	30.00	.72	.29	.05
	3 rd year	30.00	.83	.02	.00
LS2	1 st year	30.00	.59	.22	.04
	3 rd year	30.00	.69	.01	.00
NDW	1 st year	30.00	25.53	1.04	.19
	3 rd year	30.00	25.77	1.01	.18
NDWZ	1 st year	30.00	18.23	1.25	.23
	3 rd year	30.00	18.23	1.45	.27
NDWERZ	1 st year	30.00	18.40	.48	.09
	3 rd year	30.00	18.43	.74	.14
NDWESZ	1 st year	30.00	18.21	.58	.11
	3 rd year	30.00	18.30	.60	.11
TTR	1 st year	30.00	.05	.01	.00
	3 rd year	30.00	.03	.01	.00
MSTRR	1 st year	30.00	.37	.01	.00
	3 rd year	30.00	.36	.01	.00
CTTR	1 st year	30.00	.78	.11	.02
	3 rd year	30.00	.65	.11	.02
RTRR	1 st year	30.00	1.10	.15	.03
	3 rd year	30.00	.93	.15	.03
LOGTTR	1 st year	30.00	.52	.02	.00
	3 rd year	30.00	.49	.02	.00
UBER	1 st year	30.00	5.64	.08	.02
	3 rd year	30.00	5.66	.07	.01
VV1	1 st year	30.00	.05	.02	.00
	3 rd year	30.00	.04	.01	.00
SVV1	1 st year	30.00	.05	.02	.00
	3 rd year	30.00	.04	.01	.00

CVV1	1 st year	30.00	.16	.03	.00
	3 rd year	30.00	.13	.02	.00
LV	1 st year	30.00	.03	.01	.00
	3 rd year	30.00	.02	.01	.00
VV2	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
NV	1 st year	30.00	.02	.01	.00
	3 rd year	30.00	.01	.01	.00
ADJV	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
ADV	1 st year	30.00	.01	.00	.00
	3 rd year	30.00	.01	.00	.00
MODV	1 st year	30.00	.22	.01	.00
	3 rd year	30.00	.22	.01	.00

Inferential Statistics:

Note. The significance of the difference between the lexical richness scores of the two groups was assessed by examining the bolded values in the two-sided p column. Based on the p-value, a determination was made regarding whether the observed difference was likely due to mere chance. If the p-value was found to be below 0.05, the difference was considered statistically significant, suggesting it was unlikely to have occurred by chance. Conversely, if the p-value exceeded 0.05, the difference was interpreted as having a high probability of resulting from random variation or mere chance.

Table 4

Holistic Inferential Statistics

Independent Samples Test											
		Levene's Equality of Variances	Test for	t-test for Equality of Means							
		F	Sig.	T	df	Significance		Mean Difference	Std. Error Difference	95% Interval Difference	Confidence of the
						One-Sided p	Two-Sided p			Lower	Upper
LD	Equal variances assumed	1.22	.27	.77	58.00	.22	.45	.00	.00	.00	.01
	Equal variances not assumed			.77	54.99	.22	.45	.00	.00	.00	.01
LS1	Equal variances assumed	20.15	<.001	-1.96	58.00	.03	.06	-.10	.05	-.21	.00
	Equal variances not assumed			-1.96	29.38	.03	.06	-.10	.05	-.21	.00
LS2	Equal variances assumed	22.26	<.001	-2.31	58.00	.01	.02	-.09	.04	-.17	-.01
	Equal variances not assumed			-2.31	29.12	.01	.03	-.09	.04	-.18	-.01
NDW	Equal variances assumed	.05	.83	-.88	58.00	.19	.38	-.23	.26	-.76	.30

	Equal variances not assumed					- .88	57.93	.19	.38	-.23	.26	-.76	.30
NDWZ	Equal variances assumed	.79	.38	.00	58.00			.50	1.00	.00	.35	-.70	.70
	Equal variances not assumed			.00	56.73			.50	1.00	.00	.35	-.70	.70
NDWERZ	Equal variances assumed	3.67	.06	-.19	58.00			.43	.85	-.03	.16	-.35	.29
	Equal variances not assumed			-.19	49.61			.43	.85	-.03	.16	-.35	.29
NDWESZ	Equal variances assumed	.10	.75	-.57	58.00			.29	.57	-.09	.15	-.39	.22
	Equal variances not assumed			-.57	57.94			.29	.57	-.09	.15	-.39	.22
TTR	Equal variances assumed	.25	.62	4.39	58.00	<.001	<.001		.01	.00	.01	.01	.02
	Equal variances not assumed			4.39	57.75	<.001	<.001		.01	.00	.01	.01	.02
MSTRR	Equal variances assumed	.52	.48	1.60	58.00	.06	.11		.00	.00	.00	.00	.01
	Equal variances not assumed			1.60	57.52	.06	.11		.00	.00	.00	.00	.01
CTTR	Equal variances assumed	1.09	.30	4.52	58.00	<.001	<.001		.13	.03	.07	.07	.18
	Equal variances not assumed			4.52	58.00	<.001	<.001		.13	.03	.07	.07	.18
RTRR	Equal variances assumed	1.19	.28	4.54	58.00	<.001	<.001		.18	.04	.10	.10	.26
	Equal variances not assumed			4.54	58.00	<.001	<.001		.18	.04	.10	.10	.26
LOGTTR	Equal variances assumed	1.22	.27	4.53	58.00	<.001	<.001		.03	.01	.02	.02	.04
	Equal variances not assumed			4.53	57.99	<.001	<.001		.03	.01	.02	.02	.04
UBER	Equal variances assumed	.29	.59	-.72	58.00	.24	.47		-.01	.02	-.05	-.05	.03
	Equal variances not assumed			-.72	55.63	.24	.47		-.01	.02	-.05	-.05	.03
VV1	Equal variances assumed	.78	.38	4.92	58.00	<.001	<.001		.02	.00	.01	.01	.03
	Equal variances not assumed			4.92	55.16	<.001	<.001		.02	.00	.01	.01	.03
SVV1	Equal variances assumed	.78	.38	4.92	58.00	<.001	<.001		.02	.00	.01	.01	.03
	Equal variances not assumed			4.92	55.16	<.001	<.001		.02	.00	.01	.01	.03
CVV1	Equal variances assumed	.09	.77	5.11	58.00	<.001	<.001		.03	.01	.02	.02	.04
	Equal variances not assumed			5.11	57.06	<.001	<.001		.03	.01	.02	.02	.04
LV	Equal variances assumed	.03	.87	4.27	58.00	<.001	<.001		.01	.00	.01	.01	.02
	Equal variances not assumed			4.27	58.00	<.001	<.001		.01	.00	.01	.01	.02
VV2	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001		.00	.00	.00	.00	.01
	Equal variances not assumed			3.49	36.90	<.001	.00		.00	.00	.00	.00	.01
NV	Equal variances assumed	.05	.83	3.00	58.00	.00	.00		.00	.00	.00	.00	.01
	Equal variances not assumed			3.00	57.57	.00	.00		.00	.00	.00	.00	.01
ADV	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001		.00	.00	.00	.00	.01
	Equal variances not assumed			3.49	36.90	<.001	.00		.00	.00	.00	.00	.01
ADV	Equal variances assumed	104.08	<.001	3.49	58.00	<.001	<.001		.00	.00	.00	.00	.01
	Equal variances			3.49	36.90	<.001	.00		.00	.00	.00	.00	.01

	not assumed			9	90						
MODV	Equal variances assumed	.49	.48	5.21	58.00	<.001	<.001	.01	.00	.00	.01
	Equal variances not assumed			5.21	56.49	<.001	<.001	.01	.00	.00	.01

Effect Size of the difference:

Note. The effect sizes in the context of this study can be interpreted as the extent to which the difference between the lexical richness of the two groups can be observed by an ordinary reader of those texts written by the two groups of students. In other words, it is about the practical significance or meaningfulness of the results in the real world, as it can indicate the extent to which this difference would be perceptible to an average reader evaluating the students' written output.

Rules for the interpretation of results:

When interpreting Cohen's **d** results in SPSS, the following is a general **rule of thumb** for assessing the size of the effect:

1. **Small effect:** $d \approx 0.2d$
2. **Medium effect:** $d \approx 0.5d$
3. **Large effect:** $d \approx 0.8d$

Negative Cohen's d: A negative value indicates that the **second group** (usually the group that is subtracted in the formula) has a **higher mean** than the first group. In other words, a negative value represents the reverse direction of the effect.

- For example, if Cohen's $d = -0.5$, it suggests a **medium-sized effect** where the **second group (third-year students)** has a higher mean (average) than the first group (first-year students).

Magnitude: The absolute value of Cohen's d (ignoring the negative sign) tells you the size of the effect. The closer the absolute value is to 0, the smaller the effect, and the larger the absolute value, the larger the effect.

Table 5

Holistic Effect Sizes

Independent Samples Effect Sizes					
		Standardizer ^a	Point Estimate (d)	95% Confidence Interval	
				Lower	Upper
LD	Cohen's d	.01	.20	-.31	.70
LS1	Cohen's d	.21	-.50	-1.02	.01
LS2	Cohen's d	.16	-.60	-1.11	-.08
NDW	Cohen's d	1.02	-.23	-.73	.28
NDWZ	Cohen's d	1.36	.00	-.51	.51
NDWERZ	Cohen's d	.62	-.05	-.55	.46
NDWESZ	Cohen's d	.59	-.15	-.65	.36

TTR	Cohen's d	.01	1.13	.58	1.68
MSTRR	Cohen's d	.01	.41	-.10	.92
CTTR	Cohen's d	.11	1.17	.61	1.71
RTRR	Cohen's d	.15	1.17	.62	1.72
LOGTT R	Cohen's d	.02	1.17	.62	1.72
UBER	Cohen's d	.08	-.19	-.69	.32
VV1	Cohen's d	.01	1.27	.71	1.82
SVV1	Cohen's d	.01	1.27	.71	1.82
CVV1	Cohen's d	.02	1.32	.76	1.88
LV	Cohen's d	.01	1.10	.56	1.64
VV2	Cohen's d	.00	.90	.37	1.43
NV	Cohen's d	.01	.77	.25	1.30
ADJV	Cohen's d	.00	.90	.37	1.43
ADV	Cohen's d	.00	.90	.37	1.43
MODV	Cohen's d	.00	1.35	.78	1.90