

Check for
updates

Research Article

Section: Literature, Linguistics & Criticism

Published in Nairobi, Kenya
by Royallite Global.

Volume 5, Issue 4, 2024

**Article Information**

Submitted: 23rd October 2024

Accepted: 15th December 2024

Published: 31st December 2024

Additional information is
available at the end of the
article<https://creativecommons.org/licenses/by/4.0/>

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

To read the paper online,
please scan this QR code**How to Cite:**

Yusuf, M., Zein, T. T., & Nurlela, N. (2024). Assessing lexical density and its relevance to the CEFR level of reading materials for non-English-major students. *Research Journal in Advanced Humanities*, 5(4). <https://doi.org/10.58256/9ex2es14>

Assessing lexical density and its relevance to the CEFR level of reading materials for non-English-major studentsMuhammad Yusuf¹, T. Thyrhaya Zein¹, Nurlela Nurlela¹¹Universitas Sumatera Utara, Medan, Indonesia

Correspondence: yusuf_my@usu.ac.id

<https://orcid.org/0000-0003-0772-4565>**Abstract**

This study assesses the quality of reading materials for non-English-major students by analysing their lexical density (LD). This is significant in order to boost the successfulness of teaching-learning process by analysing the difficulty of the reading materials. This study employs a qualitative approach and content analysis method. A total of fifteen (15) reading texts from the *English for Islamic Studies* book were analysed. The data were analysed using Readable.com. The findings reveal that fourteen out of fifteen texts have high lexical density, meaning they contain a greater number of lexical items than grammatical ones. This suggests that texts with higher number of lexical density items provides more information but may also be more challenging for learners in comparing texts with higher number of grammatical items. In terms of readability and CEFR classification, seven (7) texts are categorized to be in the level of C1, and the other 8 texts are categorized as in C2 level.

Keywords: lexical density, reading text, text difficulty

1. Introduction

1.1. *The Background of the Study*

Textbooks play a crucial role in education by making a valuable contribution to teaching and learning. Textbooks present information in a logical and structured manner (Japelj Pavešić & Cankar, 2022; Sujatna, Heriyanto, & Andri, 2021) making it easier for students to follow lessons and understand concepts more easily. Therefore, textbooks are expected to be academically sound learning materials, with the ability to summarize and integrate knowledge in specific fields (van den Ham & Heinze, 2018). Textbooks also contribute to standardization and consistency in education (Fitria, 2023; Gultom & Pintubatu, 2022). The existence of textbooks ensures that students in various schools and classes receive consistent instruction and have access to the same core content, promoting educational equality across regions.

In the context of developing English teaching materials in Indonesia, many publishers and authors collaborate to create English textbooks. However, it should be noted that the art of composing an effective English requires more than just compiling information; it demands careful consideration of text complexity (To & Mahboob, 2019). Text complexity refers to the level of difficulty and sophistication of language in a text, which directly affects reading comprehension and language acquisition. (Anggia & Habók, 2023). In addition, Halliday's theory of text complexity focuses on the textual dimension of complexity, also known as unfolding text, with the fundamental view that a written text becomes complex when it manifests an expanded language texture (Mesmer et al., 2012). By understanding text complexity from an SFL perspective, educators can design textbooks that foster language proficiency, critical thinking, and engagement in the knowledge acquisition journey (Pu et al., 2022).

In today's interconnected world, mastering reading skills is crucial for learners, as it expands their vocabulary and strengthens overall language proficiency. This is an essential skill that must be prioritized to decrease illiteracy (Gopal, Maniam, & Nallaluthan, 2023; Sandhakumarin & Tan, 2023). The existence of reading texts in textbooks is vital for improving students' reading skills. Experts suggest that English lecturers should carefully choose the reading texts as authentic reading materials to facilitate learning and reduce student anxiety. (Ramadhani, Aulawi, & Ulfa, 2023; Vuković-Stamatović & Čarapić, 2024). These texts serve as fundamental tools that support the development of various reading abilities, essential for both academic success and lifelong learning (Bahrudin, 2016; Dewantara, Artini, & Wahyuni, 2022). Engaging with diverse types of reading texts fosters critical thinking and analytical skills. Textbooks often include a variety of genres, such as narratives, informational texts, and persuasive arguments. Each type of text requires different reading strategies, encouraging students to think critically about what they read. This happens since every reading material is unique (Siregar & Purbani, 2024). They learn to identify main ideas, draw inferences, and evaluate arguments, thereby enhancing their overall reading comprehension and analytical abilities.

Students' success at the university level largely depends on their text comprehension influenced by different features characterizing the text (Ramadhani et al., 2023; Spencer et al., 2019). Learners' capacity to anticipate comprehension depends significantly on the complexity of the texts (Gopal et al., 2023). One approach to assessing text complexity is through the measurement of lexical density (LD) (Putra & Lukmana, 2017), three fundamental aspects proposed by Halliday. Lexical density is defined as statistical measures that gauge the lexical richness of texts (Abu-Rabiah, 2023; Clavel-Arroitia, Pennock-Speck, & Pennock, 2021; Jaiswal, Kumar, & Budarina, 2024). It is a term frequently used to describe the ratio of content words (nouns, verbs, adjectives, and adverbs) to the total number of words (Halliday, 2008; Johansson, 2008). In other words, it is the proportion of lexical items to the number of running words (Gultom & Pintubatu, 2022; Kembaren & Aswani, 2022) and closely related

to nominalization (Istiqomah & Basthomi, 2024). Moreover, it is a measure of how informative and understandable a text is (Ismail, Yoestara, & Jamilah, 2023). By analysing LD, it can help to determine the difficulty level or suitability of textbooks for different student levels.

The CEFR offers a streamlined method for assessing language proficiency through its levels and descriptors, emphasizing language use in real-life situations (Waluyo, Zahabi, & Ruangsung, 2024). The CEFR framework identifies five language skill areas: spoken interaction, spoken production, listening, reading, and writing. It includes six proficiency levels that specify what learners can accomplish at each stage (Alrababa'h, Habashneh, & Rababa, 2024; Samsudin, Asrori, Mahliatussikah, & Algharibeh, 2023). These levels, known as the 'global scale,' are A1, A2, B1, B2, C1, and C2. A1 signifies the lowest proficiency, while C2 represents the highest. Each level falls into one of three main categories: beginner (A1, A2), intermediate (B1, B2), and advanced (C1, C2) (Escobar-Acevedo, Guerrero-García, & Guzmán-Cabrera, 2022; Mahmudah & Musthofa, 2023; Selvarajasingam, Nair, & Wider, 2023; Uri & Abd Aziz, 2019; Velleman & van der Geest, 2014; Warnby, 2024). This framework acts as a scientific guideline for evaluating learners' linguistic competence, particularly their communicative abilities, which can be applied in schools, markets, homes, and various other contexts (Dianto, Fatoni, & Kalita, 2022).

In Indonesia, cultural and religious background differences necessitate adapting the learning context to the learners' backgrounds in specific regions. For instance, adjustments can be observed in English textbooks in schools or universities with an Islamic focus (Islamic Studies) (Darmayenti et al., 2021). English textbooks in Islamic-based schools fundamentally differ from those in other general schools (Cahyo et al., 2019). These books play a crucial role in helping students understand English by integrating religious values into language learning. Due to contextual differences, the texts in each textbook should also differ. Besides contextual differences, text complexity in teaching materials is an important aspect to consider.

There are several previous relevant studies to this research dealing with lexical density. The first is Designing a CEFR-Based Reading Worksheet done by Abdullah, Talib, & Miolo (2023), in news text (Kembaren & Aswani, 2022; Khorina & Handani, 2022), in teachers' talk (Ismail et al., 2023), interpreting text (Liu & Dou, 2023), textbook (Jeon, 2022; Ramadhan, Santihastuti, & Wahjuningsih, 2017; Rizkiani, Mahdi, & Sujatna, 2022; Sujatna et al., 2021; To, Fan, & Thomas, 2013), speech (Amelia, Sinar, & Zein, 2020; Sibagariang, 2023), students' thesis (Siregar, Hara, & Sinar, 2024), students' translation work (Shiddiq, Oktaviani, Herwiana, & Suryanovika, 2023). Previous studies using different software to analyse the lexical density analysis such as <https://www.analyzemywriting.com/>

Research on LD has proven effective in schools for assessing the structure of productive students' competence and skills (Mangnejo & Abbas, 2024; Nacua & Eslit, 2022). Moreover, the suitable text level has been demonstrated to positively influence learners' motivation and their willingness to engage with reading material (Kuswoyo, 2023). Thus, the texts contained in students' textbooks should not be written arbitrarily but should be based on academic considerations and an understanding of students. By analyzing and measuring the lexical density and its relevance to CEFR level in reading text of English for Islamic Studies textbooks, the results of this research will serve as a basis for improvement and enhancing the quality of English textbooks, especially for university-level learners. Based on the background presented in the previous section, the research problem formulation in this study is "How is the lexical density in the book "English for Islamic Studies?"

2. Method

This study uses a qualitative approach with content analysis, supported by descriptive quantification. The qualitative procedures are used to determine lexical and functional word categories and to identify

clause rank categories in the text. Quantification is used to calculate lexical density using the formulas outlined earlier previously mentioned. These analyses are conducted through the lens of Systemic Functional Linguistics. This study focuses on analysing reading texts from the ‘English for Islamic Studies’ book. The data include reading texts from this textbook. The textbook is available online in both electronic book (e-book) format and as a physical copy through various marketplaces. The data obtained are analysed using the LFS approach proposed by Halliday, which provides a powerful analytical tool and is a well-developed linguistic approach in the field of education. Fifteen (15) reading texts from the “English for Islamic Studies” book were analysed to determine the level of LD. The analysis was conducted by using a text analyser on the website readable.com.

3. Result and Discussion

The fifteen (15) text were analyzed to identify lexical density (LD). The data were divided into two sections: Total Word Count and Total Unique Words. The following text statistics were found for each of the selected short stories:

Table 1. LD Statistics of Text 1

No	Text Statistics	
1	Total Word Count	322
2	Total Unique Words	142
3	Lexical Density	44.09% (High)

From table 1, it can be seen that the text statistics of text 1 “ALL PRAISE IS DUE TO GOD”, the total word count is 3,788 words, the total unique words are 322 words with the LD of 44.09%, it means that there are 322 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text is classified as having High Lexical Density (HLD), exceeding 40%.

Table 2. LD Statistics of Text 2

No	Text Statistics	
1	Total Word Count	1,171
2	Total Unique Words	430
3	Lexical Density	36.72% (Low)

From the text statistics of text 2 “WHAT IS ISLAM”, the total word count is 1,171 words, the total unique words are 430 words with the LD of 36.72%, it means that there are 430 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text is classified as Low Lexical Density (LLD) because it had a value of less than 40%.

Table 3. LD Statistics of Text 3

No	Text Statistics	
1	Total Word Count	112
2	Total Unique Words	66
3	Lexical Density	58.92% (High)

From the text statistics of text 3 “THE QUR’AN”, the total word count is 112 words, the total unique words are 66 words with the LD of 58.92%, it means that there are 66 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 4. LD Statistics of Text 4

No	Text Statistics	
1	Total Word Count	942
2	Total Unique Words	413
3	Lexical Density	43.84% (High)

From the text statistics of text 4 “Hajj”, the total word count is 942 words, the total unique words are 413 words with the LD of 43.84%, it means that there are 413 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 5. Table LD Statistics of Text 5

No	Text Statistics	
1	Total Word Count	300
2	Total Unique Words	146
3	Lexical Density	48.66% (High)

From the text statistics of text 5 “THE CREATION OF ADAM”, the total word count is 300 words, the total unique words are 146 words with LD of 48.66%, it means that there are 146 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 6. LD Statistics of Text 6

No	Text Statistics	
1	Total Word Count	330
2	Total Unique Words	140
3	Lexical Density	42.42% (High)

From the text statistics of text 6 “AKHIRAT”, the total word count is 330 words, the total unique words are 140 words with the LD of 42.42%, it means that there are 140 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 7. LD Statistics of Text 7

No	Text Statistics	
1	Total Word Count	393
2	Total Unique Words	200
3	Lexical Density	50.89% (High)

From the text statistics of text 7 “PROPHET”, the total word count is 393 words, the total unique words are 200 words with LD of 50.89%, it means that there are 200 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text HLD, surpassing 40%.

Table 8. LD Statistics of Text 8

No	Text Statistics	
1	Total Word Count	322
2	Total Unique Words	164
3	Lexical Density	50.93% (High)

From the text statistics of text 8 “THE MOSQUE IN ISLAM”, the total word count is 322 words, the total unique words are 164 words with LD of 50.93% it means that there are 164 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 9. LD Statistics of Text 9

No	Text Statistics	
1	Total Word Count	328
2	Total Unique Words	167
3	Lexical Density	50.91% (High)

From the text statistics of text 9 “ZAKAT”, the total word count is 328 words, the total unique words are 167 words with LD of 50.91%, it means that there are 167 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 10. LD Statistics of Text 10

No	Text Statistics	
1	Total Word Count	525
2	Total Unique Words	247
3	Lexical Density	47.04% (High)

From the text statistics of text 10 “THE ESSENCE OF FASTING”, the total word count is 525 words,

the total unique words are 247 words with the LD of 47.04%, it means that there are 247 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure's formula, this text has HLD, surpassing 40%.

Table 11. LD Statistics of Text 11

No	Text Statistics	
1	Total Word Count	276
2	Total Unique Words	131
3	Lexical Density	47.46% (High)

From the text statistics of text 11 "GOD", the total word count is 276 words, the total unique words are 131 words with the LD of 47.46% it means that there are 131 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure's formula, this text has HLD, surpassing 40%.

Table 12. LD Statistics of Text 12

No	Text Statistics	
1	Total Word Count	300
2	Total Unique Words	146
3	Lexical Density	48.66% (High)

From the text statistics of text 12 "THE CREATION OF ADAM", the total word count is 300 words, the total unique words are 146 words with LD of 48.66%, it means that there are 146 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure's formula, this text has HLD, surpassing 40%.

Table 13. LD Statistics of Text 13

No	Text Statistics	
1	Total Word Count	213
2	Total Unique Words	140
3	Lexical Density	65.72% (High)

From the text statistics of text 13 "THE RELIGION OF ISLAM", the total word count is 213 words, the total unique words are 140 words with the LD of 65.72% it means that there are 140 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure's formula, this text has a HLD, surpassing 40%.

Table 14. Table LD Statistics of Text 14

No	Text Statistics	
1	Total Word Count	657
2	Total Unique Words	266
3	Lexical Density	40.48% (High)

From the text statistics of text 14 “CURRENT AFFAIRS OF MUSLIMS”, the total word count is 657 words, the total unique words are 266 words with the LD of 40.48% it means that there are 266 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 15. LD Statistics of Text 15

No	Text Statistics	
1	Total Word Count	166
2	Total Unique Words	92
3	Lexical Density	55.42% (High)

From the text statistics of text 15 “ACCOMPANYING THE JANAZAH (THE DECEASED)”, the total word count 166 words, the total unique words are 92 words with the LD of 55.42%, it means that there are 92 words in the text that appears for the first time, for the second and the third times are not counted as unique words. According to Ure’s formula, this text has HLD, surpassing 40%.

Table 16. Text analysis

Text number	Total Word Count	Total Unique Words	Lexical Density (%)	Lexical Density (Grade)	CEFR Level
1	322	142	44.09	High	C2
2	1,171	430	36.72	Low	C2
3	112	66	58.92	High	C1
4	942	413	43.84	High	C1
5	300	146	48.66	High	C1
6	330	140	42.42	High	C1
7	393	200	50.89	High	C2
8	322	164	50.93	High	C2
9	328	167	50.91	High	C2
10	525	247	47.04	High	C2
11	276	131	47.46	High	C2
12	300	146	48.66	High	C2
13	213	140	65.72	High	C1
14	657	266	40.48	High	C1
15	166	92	55.42	High	C1

As shown in table 16, fourteen out of fifteen texts have HLD. It indicates a high number of lexical items in a clause. Conversely, LD indicates a relatively low number of lexical items in a clause. In other words, the higher the LD index of a text, the denser the information it provides, making the text more complex. Therefore, an increase in the LD index in certain sets of texts indicates an intensification of complexity among those sets of texts. Studies found that high lexical density (>40%) in many textbooks, potentially challenging for students (Waruwu et al., 2019; Rizkiani et al., 2022).

HLD poses difficulties for students in understanding the reading. Regarding writing, LD is a simple way to measure how informative a text is (Didau, 2013). As a general rule, texts with many lexical words tend to be specialized academic texts that can only be understood by well-educated individuals in specific fields. This value ranges from 0 to 100. Generally, the higher LD of a text, the heavier its content, and the more “unpacking” is required to understand it, while texts with LLD are easier to understand. In other words, a low number of lexical words results in easily understandable writing. However, if the number of lexical words is too low, the writing becomes meaningless and vague. Moreover, HLD can impact comprehension, gradual increases in difficulty may support learning (Rizkiani et al., 2022). The concept of LD, according to Johansson (2008), is related to how information is packaged; a text with a higher number of lexical items will contain more information than a text with a high number of grammatical items. As a result, texts with a high number of lexical items can be more challenging to understand due to the quality of information, and vice versa. Additionally, the number of words in a text does not affect the LD score. If a text has a complex sentence structure, there will be more grammatical items and fewer lexical items (Halliday, 2007).

Dealing with the result of CEFR level and reading text, 7 texts are categorized to be in the level of C1, and the other 8 texts are categorized as in C2 level. The Common Reference Levels of the Global Scale were proposed by the Council of Europe as the foundation of the CEFR. At six levels of proficiency, from “basic user” (A1, A2) through “independent user” (B1, B2) to “proficient user” (C1, C2), it is extensively described what a language user can “do” (Fitria, 2023; Sandhakumarin & Tan, 2023). It means those the texts are addressed for proficient user.

A good textbook is one of the major to contribute in the successfulness of education system (Sujatna et al., 2021). One type of analysis that helps them provide appropriate learning materials to students is through determining the texts’ lexical density (Rizkiani et al., 2022). The ability of the students to comprehend the text depends on the readability of the text. The students’ readability level is determined by the complexity of the text provided by the teachers (Ramadhani et al., 2023). A text with a lower lexical density index has a higher readability level, indicating that the text can be easily understood (Ramadhani et al., 2023). According to CEFR standards, texts with C1 and C2 levels on CEFR focus on advanced university and postgraduate courses in humanities and sciences, such as engineering, physics, astronomy, phycology, and computer science (Ramadhani et al., 2023). The higher frequency of lexical items (noun, verb, adjective, and adverb) that occur in a text, the more difficult it is to read (Halliday & Matthiessen, 2014). The findings are also relevant that most of texts have HLD and categorized for C1 and C2 levels.

5. Conclusion and Suggestions

Fourteen out of fifteen reading texts are classified as having high lexical density because the number of lexical items is higher than grammatical items. This indicates that texts with a high number of lexical items contain more information but are more challenging compared to texts with a high number of grammatical items. In relation to readability and CEFR level, 7 (seven) texts are categorized to be in the level of C1, and the other 8 texts are categorized as in C2 level.

In relevance to the conclusion, it is suggested that future research could focus on analyzing how students at different CEFR levels cope with texts categorized as C1 and C2, particularly for non-English-major students who may have limited exposure to complex texts. Moreover, research can be extended to develop reading materials with an appropriate balance between lexical and grammatical items to enhance readability and support language acquisition. Such studies could offer practical insights for educators and material developers in designing texts that align with students' proficiency levels while promoting better comprehension and engagement.

Acknowledgement

The authors would like to thank to Research Institute of Universitas Sumatera Utara for the Talenta Research Grant under the Scheme "Penelitian Terapan" year of 2023. We also thank Ms. Sheryll Rania Faradillah Syauki as our research assistant.

References

- Abdullah, S. S. T., Talib, R. R., & Miolo, S. T. (2023). Designing a CEFR-based reading worksheet for 7th grade students at SMP Muhammadiyah Tilango, Indonesia. *International Journal of Research and Review*, 10(9), 187–199. <https://doi.org/10.52403/ijrr.20230921>
- Abu-Rabiah, E. (2023). Evaluating L2 vocabulary development features using lexical density and lexical diversity measures. *LLT Journal: A Journal on Language and Language Teaching*, 26(1), 168–182. <https://doi.org/10.24071/llt.v26i1.5841>
- Alrababa'h, I. H., Habashneh, Q. Y., & Rababa, I. A. (2024). Assessing reading texts for non-native Arabic speaking students at the University of Jordan in light of the Common European Framework of Reference for languages from the students' perspective. *Theory and Practice in Language Studies*, 14(6), 1818–1827. <https://doi.org/10.17507/tpls.1406.23>
- Amelia, P., Sinar, T. S., & Zein, T. T. (2020). Lexical density and grammatical intricacy in debaters' speeches. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 4(1), 168–184. <https://doi.org/10.30743/ll.v4i1.2519>
- Anggia, H., & Habók, A. (2023). Textual complexity adjustments to the English reading comprehension test for undergraduate EFL students. *Heliyon*, 9(1), e12891. <https://doi.org/10.1016/j.heliyon.2023.e12891>
- Bahrudin, D. V. Y. (2016). The effect of textbook readability on students' reading comprehension. *Wacana Didaktika*, 4(1), 42–54. <https://doi.org/10.31102/wacanadidaktika.4.1.42-54>
- Cahyo, S. D., Muslim, M. R. U., Rahman, A. N., & Pratolo, B. W. (2019). Needs analysis of Islamic-based English reading material for the Muhammadiyah junior high school. *International Journal of Evaluation and Research in Education (IJERE)*, 8(2), 286–292. <https://doi.org/10.11591/ijere.v8i2.18647>
- Clavel-Arroitia, B., Pennock-Speck, B., & Pennock, B. (2021). Analysing lexical density, diversity, and sophistication in written and spoken telecollaborative exchanges. *Computer Assisted Language Learning Electronic Journal (CALL-EJ)*, 22(3), 230–250.
- Dewantara, K. A. K., Artini, L. P., & Wahyuni, L. G. E. (2022). Reading related activities in English textbook and how the texts are exploited in the classroom. *Journal of Education Research and Evaluation*, 6(3), 439–448. <https://doi.org/10.23887/jere.v6i3.48583>
- Dianto, T. N., P. N. J., Fatoni, A., & Kalita, S. (2022). CEFR-based beginner Arabic reading and writing curriculum design in Indonesia. *Ijaz Arabi Journal of Arabic Learning*, 5(3), 718–738. <https://doi.org/10.18860/ijazarabi.v5i3.16684>
- Escobar-Acevedo, A., Guerrero-García, J., & Guzmán-Cabrera, R. (2022). A model text recommendation system for engaging English language learners: Facilitating selections on CEFR. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 14(3), 1–8. <https://doi.org/10.21659/rupkatha.v14n3.17>
- Fitria, T. N. (2023). Common European framework of reference for languages (CEFR) standard in Bahasa Indonesia bagi penutur asing (BIPA) textbook: A competency mapping analysis. *Lingua Didaktika: Jurnal Bahasa Dan Pembelajaran Bahasa*, 17(2), 139–156. <https://doi.org/10.24036/ld.v17i2.124244>
- Gopal, R., Maniam, M., & Nallaluthan, K. (2023). Measuring Malaysian lower secondary learners' reading ability on a CEFR-aligned text. *Studies in English Language and Education*, 10(2), 614–627. <https://doi.org/10.24815/siele.v10i2.28013>
- Gultom, J. J., & Pintubatu, E. (2022). Grammatical intricacy and lexical density of reading texts of English textbook for senior high school grade XII. *BAHAS*, 33(2), 173–182. <https://doi.org/10.24114/bhs.v33i2.35972>
- Halliday, M. A. K. (2008). *Complementarities in language*. Beijing: The Commercial Press.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Introduction to functional grammar* (4th edition).

Oxon: Routledge.

- Ismail, N. M., Yoestara, M., & Jamilah, S. (2023). Comparing lexical density in teacher talks: Elementary school and higher education level. *LLT Journal: A Journal on Language and Language Teaching*, 26(1), 142–152. <https://doi.org/10.24071/llt.v26i1.4971>
- Istiqomah, F., & Basthomi, Y. (2024). Exploring nominalization and lexical density deployed within research article abstracts: A grammatical metaphor analysis. *Englisia: Journal of Language, Education, and Humanities*, 11(2), 14–28. <https://doi.org/10.22373/ej.v11i2.20390>
- Jaiswal, A., Kumar, P., & Budarina, A. O. (2024). A comparative study of lexical density and stylistic deviation in the poems of Jayanta Mahapatra and Kamala Das. *Journal of Siberian Federal University. Humanities & Social Sciences*, 17(1), 190–196.
- Jeon, J. (2022). Vocabulary analysis of CEFR level textbooks and Korean elementary English textbooks. *The Korea Association of Primary English Education*, 28(4), 103–125. <https://doi.org/10.25231/pee.2022.28.4.103>
- Johansson. (2008). *Lexical diversity and lexical density in speech and writing*. Lund: Lund University Press.
- Kembaren, F. R., & Aswani, A. N. (2022). Exploring lexical density in the New York Times. *ELLITE: Journal of English Language, Literature, and Teaching*, 7(2), 109–119. <https://doi.org/10.32528/ellite.v7i2.8795>
- Khorina, M., & Handani, W. Z. (2022). Comparing the lexical density of spoken and written news. *Linguistics and English Language Teaching Journal*, 10(2), 18–24.
- Liu, Z., & Dou, J. (2023). Lexical density, lexical diversity, and lexical sophistication in simultaneously interpreted texts: A cognitive perspective. *Frontiers in Psychology*, 14, 1276705. <https://doi.org/10.3389/fpsyg.2023.1276705>
- Mahmudah, U., & Musthofa, T. (2023). Reading skills learning in the “Arabic-Online.net” application by Saudi Electronic University based on the Common European Framework of Reference for Languages (CEFR). *Scaffolding: Jurnal Pendidikan Islam Dan Multikulturalisme*, 5(3), 370–385. <https://doi.org/10.37680/scaffolding.v5i3.3377>
- Mangnejo, T. H., & Abbas, A. (2024). Analysing the functions of lexical bundles for teaching academic writing to graduate students. *Sindh Journal of Linguistics*, 3(1), 121–134. <https://doi.org/10.58921/sjl.v3i1.22>
- Nacua, D. K. N., & Eslit, E. R. (2022). *Analysis on lexical density and readability in narrative writings: Basis for English writing program* (Dissertation, St. Michael’s College). St. Michael’s College. Retrieved from <https://rgdoi.net/10.13140/RG.2.2.29453.00481>
- Putra, D. A., & Lukmana, I. (2017). Text Complexity in Senior High School English Textbooks: A Systemic Functional Perspective. *Indonesian Journal of Applied Linguistics*, 7(2), 436–444. <https://doi.org/10.17509/ijal.v7i2.8352>
- Ramadhan, G., Santihastuti, A., & Wahjuningsih, E. (2017). The lexical density of Erlangga’s “English On Sky” analysis. *Jurnal Edukasi*, 4(3), 38–40. <https://doi.org/10.19184/jukasi.v4i3.6299>
- Ramadhani, R., Aulawi, H., & Ulfa, R. L. (2023). Readability of reading texts as authentic materials issued by ChatGPT: A systemic functional perspective. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 8(2), 149. <https://doi.org/10.21093/ijeltal.v8i2.1546>
- Rizkiani, D., Mahdi, S., & Sujatna, E. T. S. (2022). Lexical density and readability of the Facil’s ‘Advanced Learning Textbook’ for Indonesian high school students. *AL-ISHLAH: Jurnal Pendidikan*, 14(1), 741–752. <https://doi.org/10.35445/alishlah.v14i1.1157>
- Samsudin, M. Z., Afifarrasyihab R., Asrori, I., Mahliatussikah, H., & Algharibeh, A. A. M. (2023). Al-’Arabiyyah Baina Yadaik Book speaking skill material based on Common European Framework of Reference (CEFR). *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*,

- 6(1), 13–30. <https://doi.org/10.22219/jiz.v6i1.23141>
- Sandhakumarin, Y., & Tan, K. H. (2023). ESL teachers' views on the implementation of CEFR in reading skill. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 8(4), e002275. <https://doi.org/10.47405/mjssh.v8i4.2275>
- Selvarajasingam, S., Nair, S. M., & Wider, W. (2023). The effects of bubble map and tree map method in CEFR reading comprehension. *World Journal of English Language*, 13(7), 307–314. <https://doi.org/10.5430/wjel.v13n7p307>
- Shiddiq, S. M., Oktaviani, L., Herwiana, S., & Suryanovika, C. (2023). Lexical and grammatical errors in Indonesian-English translated texts: A text analysis on Indonesian EFL students' translation work. *JEES (Journal of English Educators Society)*, 8(1), 104–113.
- Sibagariang, S. (2023). Lexical density and readability of Joe Biden's victory speech. *Mimbar Ilmu*, 28(3), 462–468. <https://doi.org/10.23887/mi.v28i3.68448>
- Siregar, M. S., Hara, N. T., & Sinar, T. S. (2024). Lexical density and readability in students' thesis introductions. *KLAUSA (Kajian Linguistik, Pembelajaran Bahasa, Dan Sastra)*, 8(1), 71–79. <https://doi.org/10.33479/klausa.v8i1.948>
- Siregar, T. M., & Purbani, W. (2024). Prominent linguistic features of pedagogical texts to provide consideration for authentic text simplification. *Studies in English Language and Education*, 11(1), 321–342. <https://doi.org/10.24815/siele.v11i1.30815>
- Spencer, M., Gilmour, A. F., Miller, A. C., Emerson, A. M., Saha, N. M., & Cutting, L. E. (2019). Understanding the influence of text complexity and question type on reading outcomes. *Reading and Writing*, 32(3), 603–637. <https://doi.org/10.1007/s11145-018-9883-0>
- Sujatna, E. T. S., Heriyanto, H., & Andri, S. (2021). Lexical density and variation in Indonesian folklores in English student textbooks: An SFL study. *Leksika: Jurnal Bahasa, Sastra Dan Pengajarannya*, 15(2), 62–67. <https://doi.org/10.30595/lks.v15i2.11102>
- To, V., Fan, S., & Thomas, D. (2013). Lexical Density and Readability: A Case Study of English Textbooks. *Internet Journal of Language, Culture and Society*, (37), 60–71.
- Uri, N. F. M. U., & Abd Aziz, M. S. (2019). Teachers' evaluation of the suitability of reading syllabus specifications to the CEFR. *Journal of English Language Teaching and Applied Linguistics*, 1(1), 32–44.
- Velleman, E., & van der Geest, T. (2014). Online test tool to determine the CEFR reading comprehension level of text. *Procedia Computer Science*, 27, 350–358. <https://doi.org/10.1016/j.procs.2014.02.039>
- Vuković-Stamatović, M., & Čarapić, D. (2024). Vocabulary profile, lexical density and speech rate of science podcasts: How adequate are they for EAP and ESP listening? *Ibérica*, (47), 201–226. <https://doi.org/10.17398/2340-2784.47.201>
- Waluyo, B., Zahabi, A., & Ruangsung, L. (2024). Language assessment at a Thai university: A CEFR-based test of English proficiency development. *REFlections*, 31(1), 25–47. <https://doi.org/10.61508/refl.v31i1.270418>
- Warnby, M. (2024). Relating academic reading with academic vocabulary and general English proficiency to assess standards of students' university-preparedness – the case of IELTS and CEFR B2. *Scandinavian Journal of Educational Research*, 1–18. <https://doi.org/10.1080/00313831.2024.2318434>
- Waruwu, N. R., Sinulingga, J., & Erlita, Y. (2019). Lexical density and readability in English for Junior High School VIII textbook. *GENRE Journal of Applied Linguistics of FBS Unimed*, 7(4), 1–14. <https://doi.org/10.24114/genre.v7i4.12451>