LingPoet: Journal of Linguistics and Literary Research

Lexical Density in Children's Story Book

Ade Novira Aswani[1], Alemina Br. Perangin-angin[2], Aprilza Aswani[3]*

[1,2,]*Universitas Sumatera Utara, Medan, Indonesia*
[3]*Politeknik Negeri Medan, Medan, Indonesia*

Abstract. The purpose of this study was to examine the lexical density of six children's stories in order to assess the richness and complexity of language usage in these texts. The study examined the selected stories from a children's storybook using a qualitative content analysis technique to identify the fraction of content words, such as nouns, verbs, adjectives, and adverbs, that contribute to the total lexical density. To measure the lexical density of the storybooks, the analytical approaches were inspired by Ure (1971) and Halliday (1985). The findings demonstrated varying degrees of lexical density in children's storybooks, providing insights into the variable vocabulary and linguistic complexity used in these writings.

1   Introduction

Reading storybooks together has become a common practice for parents and children at home, fostering a love for reading and language (Chuang, 2023). This shared activity nurtures imagination, builds vocabulary, and strengthens the parent-child bond while creating a foundation for lifelong learning. According to Scholastic's Kids and Family Reading Report, over 50% of the children surveyed engaged in reading for enjoyment and entertainment. However, storybook reading offers much more than just entertainment value; it holds significant educational benefits. By implementing basic literacy practices while reading with their children, parents can greatly enhance the literacy experience, broaden children's comprehension, and encourage reflective thinking during the reading process.

Storybooks have a significant impact on children's cognitive and social development, offering them opportunities to acquire scientific knowledge and enhance social skills. Research consistently shows that storybook reading expands children's vocabulary and deepens their understanding of word meanings (Flack et al., 2018). Furthermore, storybooks provide a

---

*Corresponding author at: Universitas Sumatera Utara, Medan, Indonesia

E-mail address: alemina@usu.ac.id

platform for children to learn appropriate social behavior, helping them navigate social interactions and build positive relationships (Larsen et al., 2018). Storybooks also introduce children to new concepts, stimulating their curiosity and expanding their knowledge base (Gripshover & Markman, 2013). Overall, storybooks serve as powerful tools for language development, social-emotional growth, and intellectual enrichment in children's lives.

Reading is of utmost importance in a child's development as it enhances language skills, fosters imagination, and promotes cognitive growth. Children's storybooks like *Halo Balita (Hello Kids)* serve as a valuable tool in this process. Through reading such books, children not only expand their vocabulary and language comprehension but also embark on imaginative journeys filled with captivating characters and engaging narratives. *Halo Balita* specifically contributes to their linguistic and cognitive development by introducing new words, encouraging active engagement, and stimulating critical thinking. By immersing themselves in the pages of *Halo Balita*, children can embark on a delightful reading experience that nurtures their love for books, enhances their language abilities, and fosters their overall cognitive growth.

Halo Balita (Hello Kids), published by Pelangi Mizan, is a book that encompasses three distinct sections: self-help, value, and spiritual. In the self-help section, children are encouraged to cultivate independence and develop essential life skills. The Value section is dedicated to imparting moral values and principles through engaging stories and relatable examples. The Spiritual section delves into topics related to faith, spirituality, and personal growth, providing young readers with insights into their inner selves and their connection to the world around them. Each section offers unique content and valuable lessons, creating a comprehensive and enriching reading experience for children.

Taken from various sources, here are several characteristics of a good children's story book, namely: (1) compelling plot, (2) memorable characters, (3) age-appropriate language, (4) clear and concise writing, and (5) simple sentence structure. Points (3), (4), and (5) are related to lexical density. The use of simple language and sentence structures in children's storybooks leads to a lower lexical density, allowing young readers to easily understand and engage with the text. This approach fosters a positive reading experience and promotes the development of children's literacy skills. By using familiar words and straightforward sentences, children's storybooks prioritize comprehension and encourage a love for reading.

Lexical density refers to the extent of information conveyed using words in a text. It is measured by the percentage of content words, such as nouns, verbs, adjectives, and adverbs, in relation to the total number of words in the text (Halliday & Martin, 1985). According to Ure (1971), spoken texts typically have a lexical density below 40%, whereas written texts generally have a lexical density of 40% or higher. Children's storybooks, known for their simplicity and accessibility, often exhibit a lower lexical density. This is because they utilize a limited vocabulary and employ simple sentence structures that are suitable for young readers. The

intentional use of lower lexical density in children's storybooks ensures that young readers can easily understand and engage with the text, promoting language development and fostering a positive reading experience. Therefore, for children's storybooks the lexical density must be around 40% or even lower.

The lexical density of a text can be various based on the ratio of lexical items to grammatical ones, impacting the level of comprehension and difficulty for the reader (Bani Amer, 2021; Rizkiani et al., 2022). Higher lexical density, characterized by a greater proportion of lexical items, can make a text more challenging to understand, while also containing a wealth of information (Gerot, 1994; Halliday & Martin, 1985). Lexical density, obtained by dividing the number of content words by the total word count, reflects the richness and complexity of vocabulary within a text (Sari, 2016). Children's storybooks can be analyzed for their lexical density to understand the depth of vocabulary and information they offer in a concise format, providing valuable insights into language complexity and its influence on children's vocabulary acquisition.

During the analysis, the researchers made use of Analyze My Writing (https://www.analyzemywriting.com/), an online text content analyzer, to examine the lexical density of the selected stories from Halo Balita (Hello Kids). This tool enabled them to determine the percentage of different parts of speech within the texts. Additionally, the researchers accessed the online Merriam-Webster Dictionary (https://www.merriam-webster.com/), an English-to-English resource, to gain further insights into the definitions and categorization of words. By utilizing this dictionary, they were able to identify the specific part of speech for each lexical item encountered in the children's storybooks, deepening their understanding of the linguistic characteristics and vocabulary usage within the texts.

Prawinanto and Bram (2020) investigated 116 sentences containing adjective and noun clause to identify the lexical density in an English textbook for Senior High School. The study aimed to determine whether the textbook's lexical density was appropriate for the pupils. To estimate the average level of the lexical density measurements, the researchers used the lexical density theory and the lexical density indicators.

Noviyenti (2021) analyzed 16 passages from English Alive Textbook by using a content analysis technique, the textbook study focused on two domains: genres and lexical density. The study revealed that English Alive textbook contained five genres: narrative, report, analytical exposition, parody, and hortatory exposition. Based on the findings of the lexical density analysis, most of the passages are classified as readily comprehensible.

Amer (2021) examined the total of 18 texts of the reading sections in secondary stage English textbooks for 12th grade in Jordan and 20 texts of the reading sections in secondary stage English textbooks for 11th grade in Jordan. The study discovered that the English textbooks for

11th grade had a greater lexical density and readability than the 12th grade. It was also discovered that the average lexical density and readability in each grade displayed a changing pattern.

Rizkiani, Mahdi and Sujatna (2022) examined 11 reading texts selected from each chapter in the 'Advanced Learning English 2' textbook for their study on lexical density and readability levels. The results show that, except for Text 2, practically all texts had high lexical density. If the average amount is computed while including the results of the second text, the average result is 55.8%; if the average amount is calculated without including the results of text 2, the average number of lexical densities is 58.1%.

Fadhil, Gunawan and Wirza (2023) conducted a comparative analysis to identify the disparity in lexical density in English reading materials, particularly in Indonesian EFL textbooks, between government and non-government publishers. In order to discover the lexical density in English reading materials from government and non-government publishers, they used the quantitative design in this study to assess how dense the lexical elements are in these two EFL textbooks.

This study aims to examine the suitability of the children's storybook Halo Balita (Hello Kids) for children under the age of 5 by analyzing its lexical density. The focus is on six selected stories from the book, which is divided into three sections: self-help, spiritual, and value, with two stories representing each section. While existing research has predominantly explored the lexical density of English textbooks for EFL students, limited attention has been given to the lexical density of children's storybooks. Therefore, this study seeks to determine if the lexical density of the book aligns with the cognitive abilities and language development of young children. The findings will provide valuable insights into the linguistic characteristics of these stories and their potential impact on early childhood literacy.

2   Method

In this research, a qualitative content analysis approach was employed to assess the qualitative data, specifically focusing on the lexical density. The primary data for this study were the six chosen stories from the children's storybook Halo Balita (Hello Kids), published by Mizan, with two stories selected from each section. These stories were analyzed to examine the lexical density in the context of children's storybooks.

The lexical density of the six chosen stories from the children's storybook Halo Balita (Hello Kids), published by Pelangi Mizan, was analyzed using the online text analyzer, Analyze My Writing (http://www.analyzemywriting.com) which helped the researchers to find out the lexical density of the story samples. Additionally, the online Merriam-Webster Dictionary (https://www.merriam-webster.com/) was utilized to determine the part of speech for each word

in the sentence samples. The calculation of lexical density was based on Halliday's (1985) technique, providing insights into the linguistic characteristics of these stories.

3   Result and Discussion

The children's storybook Halo Balita (Hello Kids), which were firstly published in 2005 by Pelangi Mizan, consisted of three sections, namely self-help, value, and spiritual. There were nine stories in self-help section, eleven stories in value section, and five stories in spiritual section. Six stories were chosen to be the samples of the study, as described in the methodology section. Table 1 listed the titles and sections of the six stories.

Table 1 Titles and Sections

No	Titles	Sections
1	I dare to sleep alone	Self-help
2	I love books	Self-help
3	I learn how to throw out the garbage	Value
4	I love my family	Value
5	I learn how to fast	Spiritual
6	I am beautiful wearing a jilbab	Spiritual

a.   Lexical Density Analysis

This part addressed the research question: How is the realization of lexical density percentage portrayed in children's storybook Halo Balita (Hello Kids)?

Table 2 displayed the proportion of parts of speech present in the whole articles, as well as the results of the lexical density analysis.

Table 2 Lexical Density Analysis

%	Story 1	Story 2	Story 3	Story 4	Story 5	Story 6
Nouns	32.91%	30.6%	37.11%	34.84%	25.49%	27.12%
Adjectives	7.26%	4.48%	2.06%	5.88%	5.49%	5.37%
Verbs	17.09%	19.03%	19.07%	11.31%	22.75%	19.49%
Adverbs	7.69%	7.09%	4.64%	4.98%	10.2%	6.21%
Prepositions	4.27%	8.96%	7.22%	3.17%	4.71%	6.21%
Pronouns	3.85%	8.96%	9.28%	17.65%	9.41%	9.04%
Auxiliary Verbs	11.11%	4.1%	3.61%	4.07%	7.45%	6.78%
Lexical Density	**64.96%**	**61.19%**	**62.89%**	**57.01%**	**63.92%**	**58.19%**

Table 2 indicates that the lexical density of the six story samples ranged from 57.01% to 64.96%, with an average of 61.36%. This proportion was much higher than Ure's (1971) minimum percentage of lexical density in written texts. The table also found that the noun was the most regularly used component of speech in the article samples, ranging from 25% to 37% on average, providing Halo Balita (Hello Kids) storybook with a high proportion of lexical density. Table 2 displayed the proportion of parts of speech present in the whole articles, as well as the results of the lexical density analysis.

b. Data Display

The following sentences in six chosen story samples used a high percentage of content words. The chosen sentences for each text written below were the ones which had the highest and the lowest percentage of content words.

1. **"Wow, awesome!" cried Sali happily.** [Story 1, lexical density = 100% and the lexical items used in this sentence were nouns (Wow, Sali), adjectives (awesome), verb (cried), adverb (happily)].

2. "There's no **need** to be **afraid**…." [Story 1, lexical density = 28.57% and the lexical items used in this sentence were noun (need) and adjectives (afraid)].

3. **Sali also loved reading books.** [Story 2, lexical density = 100% and the lexical items used in this sentence were nouns (Sali, books), adverb (also), verbs (loved, reading)].

4. "Hey, and an **elephant too**!" [Story 2, lexical density = 40% and the lexical items used in this sentence were noun (elephant) and adverb (too)].

5. **"Tongki, caaatttccchhh...!" cried Sali.** [Story 3, lexical density = 100% and the lexical items used in this sentence were nouns (Tongki, Sali) and verbs (catch, cried)].

6. "**Sali**, I've **got** a **friend** for you." [Story 3, lexical density = 42.86% and the lexical items used in this sentence were nouns (Sali, friend) and verb (got)].

7. **Look! Even Kumi jumped** up **joyfully**. [Story 4, lexical density = 90% and the lexical items used in this sentence were verbs (look, jumped), adverb (even, joyfully), noun (Kumi)].

8. I **love** them all! [Story 4, lexical density = 25% and the only lexical item used in this sentence was verb (love)].

9. **"Sure, Dear," answered Mother.** [Story 5, lexical density = 100% and the lexical items used in this sentence were adjective (sure), nouns (dear, Mother), verb (answered)].

10. "If we can **control** ourselves, it will be **easy** to be **kind** to **others**." [Story 5, lexical density = 28.57% and the lexical items used in this sentence were verb (control), adjectives (easy, kind), noun (others)].

11. **Saliha nodded.** [Story 6, lexical density = 100% and the lexical items used in this sentence were noun (Saliha) and verb (nodded)].

12. "And when you **bend** your **body**, your **hair** won't **block** your **sight**." [Story 6, lexical density = 41.67% and the lexical items used in this sentence were verbs (bend, block) and nouns (body, hair, sight)].

Table 3 summarized the findings from the examination of the six stories under the three different sections, stories 1 up to 6. Here was the conclusion:

Table 3 Lexical Density Analysis of the Six Stories

No	Titles	Sections	Lexical Density
1	I dare to sleep alone	Self-help	64.96%
2	I love books	Self-help	61.19%
3	I learn how to throw out the garbage	Value	62.89%
4	I love my family	Value	57.01%
5	I learn how to fast	Spiritual	63.92%
6	I am beautiful wearing a jilbab	Spiritual	58.19%

Each story title has a distinct lexical density percentage. According to Ure's (1985) lexical density hypothesis, a text is classified as written language if its lexical density surpasses 40%. According to the research, all six stories of the children's book Halo Balita (Hello Kids) had lexical density percentages that exceeded 40%; all of them were even higher than 55%, suggesting that they were all eligible for written texts. The research on the six chosen stories demonstrated that the higher the lexical density percentage of a text, the more information it might express. But, remembering that this book is addressed to 5 year-old kids and younger, the content of the book is considered very complex.

The findings revealed that qualitative analysis of written texts—in this instance, the children's storybook—using online software looked to be a worthwhile endeavor. The websites were quite useful to the researcher in reviewing the data and filling out the tables. This technology may be able to distinguish between written texts created by experts and those that are not. Utilizing such techniques to analyze written texts may benefit readers, in this case parents who choose reading materials for their children, as they may determine whether Halo Balita (Hello Kids) is one of the most suitable sources for learning reading skill at an early age. Further study, however, was required to evaluate how reliable the online text analyzer would be in analyzing the written texts.

4   Conclusion

Based on the findings, it is possible to infer that Halo Balita (Hello Kids) can be considered as a complex written text for a children's storybook because the six story samples had a high lexical density percentage. It is based on Ure and Halliday's ideas about lexical density; the written text must contain at least 40% lexical density, which means this book contains many information and vocabularies for children to learn, but it can be too much for younger kids to grasp the content. Therefore, parents must assist their children while reading this book. The lexical

density percentages for the story samples are 64.96% for story 1, 61.19% for article 2, 62.89% for article 3, 57.01% for article 4, 63.92% for article 5, and 58.19% for article 6.

REFERENCES

[1] Bani Amer, M. (2021). *Lexical Density and Readability of Secondary Stage English Textbooks in Jordan*. *2*, 11–20.

[2] Chuang, J. (2023). Children's Textual Transactions In Parent-Child Storybook Reading At Home. *Journal of Classroom Research in Literacy*, *13*. https://jcrl.library.utoronto.ca/index.php/jcrl/article/view/38447

[3] Fadhil, A., Gunawan, W., & Wirza, Y. (2023). Lexical Density in EFL Indonesian Textbooks: A Comparative Analysis. *JALL (Journal of Applied Linguistics and Literacy)*, *7*(1), Article 1. https://doi.org/10.25157/jall.v7i1.9727

[4] Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/10.1037/dev0000512

[5] Gerot. (1994). *Making Sense Of Functional Grammar (Gerot – Wignell)*.

[6] Gripshover, S. J., & Markman, E. M. (2013). Teaching Young Children a Theory of Nutrition: Conceptual Change and the Potential for Increased Vegetable Consumption. *Psychological Science*, *24*(8), 1541–1553.

[7] Halliday, & Martin. (1985). *HALLIDAY, M. A. K.; MARTIN, J. R. (eds.) Writing Science: Literacy and Discursive Power. Bristol/London: The Falmer Press, 1993.* https://www.academia.edu/21703203/HALLIDAY_M_A_K_MARTIN_J_R_eds_Writing_Science_Literacy_and_Discursive_Power_Bristol_London_The_Falmer_Press_1993

[8] Larsen, N. E., Lee, K., & Ganea, P. A. (2018). Do storybooks with anthropomorphized animal characters promote prosocial behaviors in young children? *Developmental Science*, *21*(3), e12590. https://doi.org/10.1111/desc.12590

[9] Noviyenti, L. (2021). Analysis of English Alive Textbook in Terms of Genres and Lexical Density. *AL-ISHLAH: Jurnal Pendidikan*, *13*(2), Article 2. https://doi.org/10.35445/alishlah.v13i2.587

[10] Prawinanto, & Bram. (2020). *Adjective And Noun Clause Lexical Density In An English Textbook For Senior High School Students | Getsempena English Education Journal*. https://ejournal.bbg.ac.id/geej/article/view/1019

[11] 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), Article 1. https://doi.org/10.35445/alishlah.v14i1.1157

[12] Sari, D. (2016). Measuring Quality Of Reading Materials In English Textbook: The Use Of Lexical Density Method In Assessing Complexity Of Reading Materials Of Indonesia's Curriculum – 13 (K13) English Textbook Dian Sari. *JOALL (Journal of Applied Linguistics and Literature)*, *1*(2), Article 2. https://doi.org/10.33369/joall.v1i2.4177

[13] Ure, J. (1971). *Lexical density and register differentiation.–Applications of Linguistics. Selected Papers of the Second International Congress of Applied Linguistics, Cambridge 1969. Ed. George Perren, John LM Trim*. London: Cambridge University Press.