Analyzing Lexical Density and Readability of Reading Texts in English Textbook "Stop Bullying Now" by Mahrukh Bashir

Nafkhatul Miskiyah[1], Taranindya Zulhi Amalia [2]
[1,2] Tadris Bahasa Inggris, Fakultas Tarbiyah, Institut Agama Islam Negeri Kudus, Indonesia

Contact:	Nafkhatul Miskiyah Taranindya Zulhi Amalia	✉	nafa.naifa12@gmail.com khasnatara@gmail.com

ABSTRACT

The complexity of the text can be determined by lexical density and readability tests. Therefore, the purposes of this research are to find out the lexical density and readability of reading texts and the relevancy of them to the grade level of reading texts. This research took a descriptive qualitative method with library research. The subject is taken from the eleven reading texts in the English textbook "Stop Bullying Now" for eleventh grade. For analyzing data, the researcher used Ure formula and Flesch Reading Ease formula. The finding showed that 8 texts with a lexical density score >50% were categorized as quite-density and 3 texts were categorized as low-density. There are 5 levels of reading and only 4 texts which correspond to the students in senior high school. Based on the average analysis score, the texts are categorized as standard level and more appropriate for students in 8th-9th grade. In this study's data, lexical density and readability do not change considerably as text levels increase.

Introduction

Reading skills are very necessary in learning activities for students nowadays. There is a lot of information that we can get by reading a text. This is in line with Manurung and Tuahman (2019) who said that reading a text is a source of knowledge for students. It is not only to get information, but also offers learning opportunities in grammar, vocabulary, and pronunciation (Harmer, 1998). Therefore, mastering reading skills is very important. Every student must master reading skills, so that they can comprehend and interpret the meaning of text easily.

One of the learning media that is often used by schools is the textbook. Kathleen Graves (2000) describes how a textbook serves as a learning and teaching tool and motivation. Textbook as media learning plays an important role in helping teachers improve students'

language skills, especially in reading skills. It is used to facilitate students' learning activities. Teachers should be able to select the best English textbook for their students. Besides the book corresponds to the current curriculum, it is also important to provide a suitable reading text for students. The term "suitable reading texts" refers to texts that are read at the same level as the students' reading ability. Finding suitable reading texts in an English textbook is not easy. For some students, the text is either too easy or too complex. The facts also reveal that many students have difficulty understanding English texts. A teacher of one senior high school in Jepara said that most students find it difficult in vocabulary and pronunciation when they must read and comprehend a text that has many paragraphs because there are lots of unknown vocabulary for them. All of these are directly related to the lexical density and readability of the texts.

Lexical density is related to the amount of information in a text that can be analyzed by counting the content words (lexical items) in a text. It predicts that text with a large number of lexical items is more difficult to comprehend than text with low lexical items because there is more information (Johansson, 2008). And readability refers to the level of ease or difficulty of a text. In other words, the degree to which what is conveyed by written text is easier or more difficult to understand is readability (Bailin & Grafstein, 2016). Therefore, lexical density and readability influence the difficulty levels of a text. Knowing the lexical density and readability of a text will help the teacher to find the appropriate reading text for their students. Some studies also prove that lexical density and readability are two of the factors that influence the complexity levels of a text. Daw Thida (2019), states that written text with a lexical density score of more than 50% is in the fairly difficult to read category. In addition, To Fan and Thomas (2013) also conducted research on the complexity of the English textbook using lexical density and readability for each text at different levels. From this interesting research, they prove that lexical density and readability can measure the complexity of a text.

From this research, we are going to analyze the English textbook for eleventh grade "Stop Bullying Now" by Mahrukh Bashir. An English teacher said this book is suitable for the current curriculum and has various reading texts and a special part for reading materials. Therefore, in this research we focus on finding the lexical density and readability of reading texts in the English textbook, and the relevance of lexical density and readability to the grade level of reading texts in the English textbook "Stop Bullying Now" for eleventh grade senior high school.

Lexical Density

The lexical density of a text is used to determine how informative it is. It measures the relationship between lexical words and grammatical words in a text (Thornbury & Slade, 2006).

The concept refers to how information is presented; a text with higher lexical words will contain more information than a text with high grammatical words (Johansson, 2008). This concept was first introduced by Ure in 1971. He describes there are two categories of lexical density; lexical item and grammatical item (Stubbs, 2002). Lexical items usually carry high information, such as verbs, nouns, adjectives, and adverbs. And grammatical items include auxiliary verbs, pronouns, prepositions, determiners, conjunctions, and numerals (Stubbs, 2002). The purpose of grammatical items is to connect the content words together in a sentence, even if they do not give information (Syarif & Putri, 2018).

The concept of lexical density by Ure uses a grammatical system, not a lexical set, where they distinguish between lexical and grammatical words in their calculations (Johansson, 2008). For example, take out. Ure counts "take" as one lexical word and "out" as one grammatical word. The formula used by Ure to measure the ratio of lexical density is the total number of lexical items in a text as a proportion of total words (Stubbs, 2002).

$$\textit{Lexical density} = \frac{\textit{Number of Lexical}}{\textit{Total Number of Words}} \times 100$$

Readability

The readability test is to measure the text level. It will be able to predict comprehension at the sentence level (McNamara, Graesser, McCarthy, & Cai, 2014). Bailin & Grafstein (2016) describe readability as the degree to which what is represented by written text is simple or difficult to comprehend. One of the tested and reliable formulas for measuring the readability of a text is the Flesch Reading Ease Formula. It is determined by the index score on sentence length and the number of syllables per word (DuBay, 2004). The scale position starts from 0-100. 0 indicates a very difficult level and 100 indicates a very easy level.

RE : (206.835 - (1.015 x ASL)) - (84.6 x ASW)

ASL (Average Sentence Length) is total of words divided by total of sentences. ASW (Average Syllable per Word) is total of syllables divided by total of words.

Table 1. Readability Level by Flesch Reading Ease Formula

Reading Ease Score	Description	Reading Grade
0 - 29	Very Difficult	College Graduate
30 - 49	Difficult	13th - 16th grade
50 - 59	Fairly Difficult	10th - 12th grade
60 - 69	Standard	8th - 9th grade
70 - 79	Fairly Easy	7th grade
80 - 89	Easy	6th grade
90 - 100	Very Easy	5th grade

Method

The research method is a scientific process of gathering data in order to solve a problem. In conducting this study, we used library research, where the researcher tries to acquire data from the literature, which includes books, journals, and other materials (Fathoni, 2006). And the approach used descriptive qualitative because it focuses on analyzing and interpreting the reading material in a textbook. In general, qualitative research describes social issues in a natural way and aims to drive the important meaning from a certain event, truth, occurrence, reality, or issue (Ary, Jacobs, Sorensen, & Razavieh, 2010).

The subject of this research is from reading texts in the English textbook "Stop Bullying Now" for the eleventh grade of Senior High School. This book was published by the Ministry of Education in 2017. There are eleven reading texts that will be analyzed. To collect data, the researcher used documentation study, which is a data collection method that involves collecting, reviewing, and analyzing documents. And for data analysis technique, she used content analysis with three steps by Miles and Huberman. They are data reduction, data display, and conclusion/verification (Miles, Huberman, & Saldana, 2014).

1. Data reduction

This process started by reading and identifying the texts. For analyzing lexical density, selecting the words of lexical items and grammatical items using part-of-speech.info. Focusing and calculating the number of lexical items in a text. For readability test, we used online-utility.org and howmanysyllables.com, then focused and calculated the number of words, sentences, and syllables of text.

2. Data display

The result of data reduction is then entered into these tables.

Table 2. Data Display of Lexical Items

No	Text	Noun	Verb	Adjective	Adverb	Total lexical items

Table 3. Data Display of Readability

No	Text	Total of words	Total of sentences	Total of syllables	ASL (total of words / total of sentences)	ASW (total of syllables / total of words)

3. Conclusion / verification

To make conclusion of lexical density, Ure Formula is used

$$Lexical\ density = \frac{Number\ of\ Lexical\ Items}{Total\ Number\ of\ Words} X\ 100$$

If the score more than 60% is high lexical density. 50-60% is quite lexical density. And less than 50% is low lexical density.

For readability test used Flesch Reading Ease Formula.

RE   : (206.835 - (1.015 x ASL)) - (84.6 x ASW)

ASL   : total of words / total of sentences

ASW   : total of syllable / total of word

Result

Summarize the collected data and the analysis performed on those data relevant to the issue that is to follow. The Findings should be clear and concise. It should be written objectively and factually, and without expressing personal opinion. The result of selecting the lexical item is displayed in the table below.

Table 4. Data Display of Lexical Items

No	Text	Noun	Verb	Adjective	Adverb	Total lexical items
1.	Global Warming	80	27	29	11	147
2.	Banning of Motorbikes	102	29	28	19	178
3.	Personal Letter	48	26	14	17	105
4.	Earthquakes	60	19	24	10	113
5.	How Volcanoes are Formed?	45	20	5	3	73
6.	The Enchanted Fish	178	173	65	61	477
7.	Bullying	164	118	66	41	389
8.	President Sukarno of Indonesia	124	53	45	30	252
9.	Letter to God	254	197	59	52	562
10.	The Last Leaf	242	160	69	68	539
11.	Ki Hajar Dewantara	223	62	51	14	350

From this table, we count the presentation of lexical density used Ure Formula. And the result can be seen in this table.

Table 5. Data Result of Lexical Density

No	Text	Lexical Items	Grammatical Items	Word	Lexical Density	Level LD
1.	Global Warming	147	125	272	54, 04	Quite
2.	Banning of Motorbike	178	147	325	54, 77	Quite
3.	Personal Letter	105	109	214	49,07	Low
4.	Earthquakes	113	83	196	57, 65	Quite
5.	How Volcanoes are Formed?	73	68	141	51, 77	Quite
6.	The Enchanted Fish	477	548	1025	46, 54	Low
7.	Bullying	389	383	772	50, 39	Quite
8.	President Sukarno	252	250	502	50, 20	Quite
9.	Letter to God	562	585	1147	48, 99	Low
10.	The Last Leaf	539	482	1021	52, 79	Quite
11.	Ki Hajar Dewantara	350	319	669	52, 32	Quite
AVERAGE					51, 68%	**Quite**

According to the table above, eight of the eleven reading texts were classified as high lexical density because the lexical items were higher than the grammatical items, while three others were classified as low lexical density because there were more grammatical items than lexical items. There were no texts classified as having high lexical density because the score was less than 60%. It indicates that text with high lexical items contains more information but is more difficult than text with high grammatical items. For a readability test, the first step is counting the number of words, sentences, and syllables in a text. The result can be seen in table 6. The next step was counting the readability of the text using the Flesch Reading Ease Formula. The results of the lexical density and readability test are combined into table 7.

The grade level of each text can be assessed by looking at the lexical density and readability scores. From the table above, there were only four texts that were suitable for students in eleventh grade of senior high school. These texts were categorized as fairly difficult level. Those texts were text 2 (Banning of Motorbike), text 4 (Earthquakes), text 7 (Bullying), and text 8 (President Sukarno of Indonesia). And the lexical density score from the four texts was at the same level, which was classified as quite dense.

Table 6. Data Display of Readability

No	Text	Total of words	Total of sentences	Total of syllables	ASL (total of words / total of sentences)	ASW (total of syllables / total of words)
1.	Global Warming	272	16	455	17, 00	1, 67
2.	Banning of Motorbike	325	21	530	15, 48	1, 63
3.	Personal Letter	214	19	306	11, 26	1, 43
4.	Earthquakes	196	12	316	16, 33	1, 61
5.	How Volcanoes are Formed?	141	8	195	17, 63	1, 38
6.	The Enchanted Fish	1025	81	1248	12, 65	1, 22
7.	Bullying	772	45	1212	17, 16	1, 57
8.	President Sukarno	502	34	788	14, 76	1, 57
9.	Letter to God	1147	92	1548	12, 47	1, 35
10.	The Last Leaf	1021	99	1338	10, 31	1, 31
11.	Ki Hajar Dewantara	669	35	1210	19, 11	1, 81

Table 7. Data Display of Grade Level of Texts

No	Text	Lexical Density	Readability	Level LD	Level Text	Grade Level
1.	GW	54, 04	48, 30	Quite	Difficult	13th-16th
2.	BoF	54, 77	53, 22	Quite	Fairly difficult	**10th-12th**
3.	PL	49,07	74, 43	Low	Fairly easy	7th
4.	EQ	57, 65	54, 05	Quite	Fairly difficult	**10th-12th**
5.	VOL	51, 77	72, 19	Quite	Fairly easy	7th
6.	TEF	46, 54	90, 78	Low	Very easy	5th
7.	BLYNG	50, 39	56, 60	Quite	Fairly difficult	**10th-12th**
8.	PSoI	50, 20	59, 03	Quite	Fairly difficult	**10th-12th**
9.	LtG	48, 99	79, 97	Low	Fairly easy	7th
10.	TLL	52, 79	85, 54	Quite	Easy	6th
11.	KHD	52, 32	34, 31	Quite	Difficult	13th-16th
Average score		**51, 68**	**64, 40**	**Quite**	**Standard**	**8th-9th**

Discussion

Lexical density of a text is used to determine how informative it is. Based on Sholichatun (2011), text with a lexical density score of 40-50% is low dense and the score of 50-60% is quite dense. From this analysis, it found that from 11 texts, there were 3 texts classified as low density and 8 texts classified as quite density. No one text was classified as high lexical density because the score was less than 60%. The texts with low density were Text 3, Text 6,

and Text 9. And the texts with quite lexical density were Text 1, Text 2, Text 4, Text 5, Text 7, Text 8, Text 10, and Text 11. There were more lexical items in these texts than grammatical items, indicating that they used a rich vocabulary to convey the text's contents.

The concept of lexical density, according to Johansson (2008), corresponds to the way information is packaged; a text with a higher lexical item will contain more information than a text with a high grammatical item. As a result, a text with a high lexical item can be more difficult to comprehend due to the quality of information, and vice versa. In addition, the number of words in a text does not affect the score of lexical density. Because the great written text features easy sentence structures and a large lexical item. If the text has a complicated sentence structure, there will be more grammatical items and fewer lexical items (Halliday, 2007).

The example can be seen in the Table 5. Text 6 (The Enchanted Fish) had 1025 words, but it was categorized as the lowest lexical density with a score of 46.54% because the grammatical item was higher than the lexical item. Meanwhile, Text 4 (Earthquakes) had 196 words but it was categorized as the highest lexical density with a score of 57.65% because it had a simple sentence structure and the lexical item was higher than the grammatical item. It indicates that the lexical density score is not affected by the number of words in a text.

To analyze the readability and the grade level of texts, we used Flesch Reading Ease formula. From this analysis, it found there were 5 levels of text.

1.   5th grade (Very Easy Level)

In the eleventh grade English textbook, there was only one text that was categorized as very easy level. This text was Text 6 (The Enchanted Fish). It had a readability score of 90,78 and was suitable for students in 5th grade because the readability score ranges between 90-100. It was predicted that students in senior high school should be able to comprehend this text without difficulty.

2.   6th Grade (Easy Level)

In the eleventh grade English textbook, there was only one text that was categorized as easy level. This text was Text 10 (The Last Leaf). It had a readability score of 85,54 and was suitable for students in 6th grade because the readability score ranges between 80-89. Students in senior high school should be able to understand this text.

3.   7th Grade (Fairly Easy Level)

In the eleventh grade English textbook, there were three texts that were categorized as fairly easy level. This text was Text 3 (Personal Letter) with readability score of 74,43. Text 5 (How Volcanoes are Formed) with a readability score of 72,19. And Text 9 (Letter to God) with a readability score of 79,97. All these texts were suitable for students in 7th grade Junior High

School because the readability score ranged from 70 to 79. Students in senior high school grade 11 were expected to comprehend these texts quite easily.

4.   10th-12th Grade (Fairly Difficult Level)

In the eleventh grade English textbook, there were four texts that were categorized as fairly difficult level. This text was Text 2 (Banning of Motorbike) with a readability score of 53,22. Text 4 (Earthquake) with readability score of 54,05. Text 7 (Bullying) with a readability score of 56,60. And Text 8 (President Sukarno) with a readability score of 59,03. All these texts were suitable and matched for students in 10th-12th grade Senior High School because the readability score ranged between 50-59.

5.   College Students

In the eleventh grade English textbook, there were two texts that were categorized as difficult level. This text was Text 1 (Global Warming) with a readability score of 47,30. And Text 11 (Life and Times of Ki Hajar Dewantara) with a readability score of 34,31. These texts were more suggested for college students because the readability score ranged between 30-39. Students in senior high school grade 11 were expected to find it difficult to comprehend these texts, but also challenging for them.

A higher readability score (on a scale of 0-100) indicates that a text is more readable (Bailin & Grafstein, 2016). It means that the text will be more difficult to comprehend if it has a lower readability score than others. The result of the lexical density and readability of the texts (Table 7) showed that the level of readability did not really meet the lexical density index. Each readability score range had a varied density score. For example, Text 3 and Text 5 that were suitable for students in 7th grade were categorized as fairly easy level. But, Text 3 was categorized as low-density and Text 5 was categorized as quite-density. As a result, higher text levels are not always associated with higher lexical density scores.

In general, the reading texts in English textbook eleventh grade had an average lexical density score of 51,68% (quite lexical density) and the average readability score was 64,40 or in the standard level category. It means these texts are, on average, categorized as quite dense and informative. For student in eleventh grade, it is predicted they could comprehend the text because it is more appropriate for 8th-9th graders in Junior High School. It can be a problem for students because the result of readability is not suitable for the grade level of students. Some texts are too easy for students in grade eleven, while others are too challenging. As a result, the teacher's participation is a key in choosing the best technique and method for assisting students in comprehending reading materials. If the material is neither readable nor appropriate for students, the teacher must be creative.

Conclusion

The following conclusions can be drawn from the findings of this study. Based on the statistical results of lexical density, 3 texts are classified as low lexical density and 8 texts are classified as quite lexical density. It means these eight texts contain more information than three texts with low lexical density. Based on the results of readability test, the eleven texts are classified into 5 reading levels. They are very easy level, easy level, fairly easy level, fairly difficult level, and difficult level. There are only 4 texts in the fairly difficult category that are appropriate for students in eleventh grade. The average lexical density score is 52,68% and the readability score is 64,40. It indicates that these reading texts are quite informative and are classified as standard level of reading. It means the reading text in this English textbook is appropriate for 8th-9th grade in Junior High School. Besides that, in this study's data, lexical density and readability do not change considerably as text levels increase.

However, the writer and publisher must consider all components contained in the textbook, including the suitability between the reading text and the student's grade level before the book is published. In addition, Teachers must always help and facilitate students in learning and understanding reading texts even though the texts are not in accordance with the grade level of students, because each student has different abilities. For future researchers who want to research the same topic, they can use this research as a reference and they can also use other types of books and modify it. So that people in the field of education can better understand in choosing the right reading text.

Acknowledgement

References

Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). *Introduction to Research in Education* (8th editio). Belmont: Wadsworth Cengage Learning.

Bailin, A., & Grafstein, A. (2016). *Readability: Text and Context* (1st editio). Macmillan: Hofstra University.

DuBay, W. H. (2004). *The Principles of Readability: A Brief Introduction to Readability Research*. *Impact Information*. California: Costa Mesa.

Fathoni, A. (2006). *Penelitian dan Tehnik Penyusunan Skripsi*. Jakarta: Rineka Cipta.

Graves, K. (2000). *Designing Language Courses: A Guide For Teacher*. (D. Freeman, Ed.). Canada: Heinle & Heinle Publisher.

Halliday, M. A. K. (2007). *Language and Education*. (J. J. Webster, Ed.). New York: Continum.

Harmer, J. (1998). *How to Teach English: An Introduction to The Practice of English Language Teaching*. (P. Education, Ed.). Essex.

Johansson, V. (2008). Lexical Diversity and Lexical Density in Speech and Writing: A Developmental Perspective. *Lund University, Dept. of Linguistic and Phonetics, Working Papers*, *53*, 61–79.

Manurung, A., & Sipayung, K. T. (2019). An Analysis LexIcal Density of Reading text in SMA Imelda Medan. *JETAL: Journal of English Teaching & Applied Linguistic*, *1*(1), 41–44. https://doi.org/10.36655/jetal.v1i1.124

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). Coh-Metrix Measures of Text Readability and Easability. In *Automated Evaluation of Text and Discourse with Coh-Metrix* (pp. 78–95). Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511894664.007

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook*. London: Sage Publication.

Sholichatun, S. (2011). *Content Analysis of Reading Materials in English on Sky Textbook for Junior High School*. UIN Walisongo.

Stubbs, M. (2002). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell Publishing.

Syarif, H., & Putri, R. E. (2018). How Lexical Density Reveals Students ' Ability In Writing Academic Text. *Lingua Didaktika*, *12*(2), 86–94. https://doi.org/10.24036/ld.v12i2.10408

Thida, D. (2019). Lexical Density and Readability of Students ' Writing. *International Journal of Science and Research (IJSR)*, *8*(9), 2018–2020.

Thornbury, S., & Slade, D. (2006). *Thornbury and Diana Slade, Conversation: From Description to Pedagogy. UK: Cambridge University Press, 2006*. Cambridge: Press, Cambridge University.

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*(37), 61–71.

This page intentionally left blank