

# Text Complexity in English Textbooks for Junior High School: *A Systemic Functional Perspective*

Wida Mulyanti, Pipit Prihartanti Soeharto  
 Universitas Perjuangan Tasikmalaya  
 Tasikmalaya, Indonesia  
 widamulyanti@unper.ac.id, pprihartanti@gmail.com

**Abstract**— This study investigates the progression of text complexity based on Systemic Functional Linguistics theory. Being largely qualitative, this study analyzed three consecutive textbooks (grade VII, VIII, and IX) issued by the Ministry of Education. The data were analyzed in terms of lexical density, lexical variation, and grammatical complexity to determine the level of complexity of the texts. The findings show that the lexical density and lexical variation fluctuated as the level progressed. It means that the lexical density and lexical variation throughout the textbooks were found to have an inconsistent development from lower to higher grades. Meanwhile, the findings in grammatical intricacy indicate that the complexity of texts decreased as the level progressed. Thus, it can be concluded that the texts required reviewing to correspond to the stages of students' intellectual development.

**Keywords:** *grammatical intricacy, lexical density, lexical variation, text complexity, Systemic Functional Linguistics, textbooks*

## I. INTRODUCTION

The textbook has been well-known to play a key role in the process of learning at school. It can determine the materials to be taught and also can help teachers decide the way they teach (see e.g. Watts-Taffe, Gwinn, Johnson & Horn, 2003; Dole & Osborn, 2003; Reys, Reys & Chávez, 2004; Apple, 1991; Elliot & Woodwart, 1990; Putra & Lukmana, 2017). According to Mukundan, Hajimohammadi, and Nimehchisalem (2011), textbook plays a crucial role in learners' success or failure in their academic program. Moreover, Palinscar and Duke (2004, p. 184) added that even when teachers prefer not to teach from the textbook, it still plays a significant role in determining the curriculum.

The importance of the textbooks mentioned above should be sufficient reason to show the importance of exploring textbooks, whether it is seen in terms of cultural, social, or textual aspects so that textbooks become effective tools for students in gaining knowledge and skills. However, Nathan, Long, and Alibali (2002) argue that although textbook is the main learning tool, the analysis of its composition and organization is often neglected in the study concerning text analysis.

Text complexity, which belongs to the textual element of the textbook, "could be considered one key element in determining the students' success in grasping the ideas contained in the textbooks" (Putra & Lukmana, 2017, p. 436).

The text complexity in a textbook should be in line with the students' development level for the texts to be comprehended well by the students so that it can stimulate their intellectual development.

Text complexity is one of the linguistic features of written texts that can affect the level of difficulty. However, even though text complexity is fundamental to the current science, its precise definition has still been a debatable issue (Kwapien, Drozd & Orczyk, 2010). Besides, Rimmer (2008) states that there is no linguistic standard in measuring text complexity.

In SFL, lexical density, lexical variation, and grammatical intricacy can be used to measure the complexity of the text. Halliday (1985a, 1987, 2014) states that written language becomes complex by being lexically dense, the density can be reduced by the repetition of lexical words, and spoken language becomes complex by being grammatically intricate.

The lexical density of texts is an important factor of complexity in written language (Halliday, 1989). Halliday and Martin (1993, p.76) define lexical density as "the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure". From SFL perspectives, lexical items include nouns, verbs, adjectives, and adverbs, while grammatical items consist of pronouns, determiners, finite verbs, and some classes of adverbs (Halliday, 1985b).

Lexical density in this study is analyzed by applying the formula proposed by Halliday. Halliday (2014) suggests that to measure lexical density, we must divide the number of lexical items by the number of ranking clauses. Halliday's lexical density (LD) formula for calculation is as follows (Castello, 2008, p. 97):

$$LD \text{ (index)} = \frac{\text{Total number of lexical words}}{\text{Total number of ranking clauses}}$$

Lexical variation is another important factor in the measurement of text complexity. Halliday (1989, p. 64) points out that "repetition reduces the effect of density". It means that "if there is an approximately equal number of grammatical and lexical items in two hypothetical texts, but in one text some lexical items are repeated several times, whereas in the other such items are used only once, then the second text is likely to be perceived as more difficult" (Putra & Lukmana, 2017).

Lexical variation (LV) formula is as follows (Castello, 2008, p. 64):

$$\text{LV (index)} = \frac{\text{Number of different lexical words}}{\text{Total number of lexical words}}$$

To make the measurement of text complexity even more accurate, the grammatical intricacy will also be measured. This quantification will allow us to get a better picture of text complexity because it measures a comparison between a clause complex appearances in a text with simple clauses. According to Halliday (2014, p.728), “the written version is more complex in terms of lexical density, while the spoken version is more complex in terms of grammatical intricacy”. Although grammatical intricacy tends to be used for the analysis of the complexity of spoken language as argued by Halliday, the measurement is still worth included in written text complexity because it can help us to know the complexity of the texts not only at the level of words but also at the level of clauses (Putra & Lukmana, 2017). The formula of Grammatical Intricacy (GI) is as follows (Castello, 2008, p. 97):

$$\text{GI (index)} = \frac{\text{Total number of ranking clauses}}{\text{Total number of clause complexes}}$$

There have been several numbers of studies regarding text complexity in the SFL theoretical framework. To (2017) focus on measuring grammatical intricacy of text. Halliday (2008) says that grammatical intricacy is an important concept in characterizing the complexity of language. The findings revealed that the grammatical intricacy increased in accordance with the book levels. The mean scores of grammatical intricacies showed a gradual increase from the elementary to the intermediate book level, though they were not significantly different, and the upper-intermediate textbook did not show the topmost grammatical complexity. To (2018) continuing his research regarding text complexity by analyzing three linguistic features namely, lexical density, nominalization, and grammatical metaphor of 24 reading texts in the chosen textbooks. The findings show that overall textbook texts grew complex when their levels advanced.

Another study regarding text complexity in the SFL framework is written by Newnham (2013). This study is the first attempt in exploring text complexity in graded readers by comparing semantically equivalent passages across different rewrites of the same title. The study analyzed the taxis, logico-semantic relations, grammatical metaphor, and appraisal structure of the texts. The findings of this study are indicative of increased structural and semantic complexity by level. These findings suggest consideration of text complexity in context, as opposed to de-contextualized word or grammar lists.

## II. METHOD

This study is mostly qualitative, supported by descriptive quantifications, and the analyses were based on the SFL perspectives. In determining the category of lexical and functional words and in determining the category of ranking clauses in the texts, the qualitative procedures were used. The quantification was used to find the indexes of lexical density, lexical variation and grammatical intricacy based on the formulas mentioned before.

This study analyzed three sequential junior high school textbooks used in Indonesian schools. The data included the reading texts from three textbooks, the book for grade VII, grade VIII, and grade IX. The textbooks selected for the study were published by the Ministry of Education of the Republic of Indonesia in 2017 and 2018 to be used as school textbooks based on the Curriculum 2013. These three textbooks are available online in electronic book format (e-book) and can be downloaded freely for classroom use.

A total of six texts were selected from the three different books and thus from three sequential grades with the following distribution: two texts from Grade VII textbook (Textbook 1), two texts from grade VIII textbook (Textbook 2), and two texts from grade IX textbook (Textbook 3). The two texts from each textbook were selected based on the chapters where the texts were situated. One text (Text 1a, Text 1b, and Text 1c) were selected from the odd semester of each textbook; other texts (Text 2a, Text 2b, and Text 2c) were selected from the even semester of each textbook. The selection was organized that way to investigate whether there is an increase of text complexity within and among textbooks.

The data were analyzed based on the Systemic Functional Linguistics theoretical framework proposed by Halliday. Freebody (2003) states that this theoretical framework is a powerful analytical tool and has been considered as one of the linguistic approaches that have been well developed in the education field. The texts which have been selected from the three textbooks were analyzed in terms of the complexity of text which covered the lexical density, lexical variation, and grammatical intricacy. The lexical density analysis and grammatical intricacy analysis were done manually. Meanwhile, the analysis of lexical variation involved an automation process by using AntConc Windows (3.5.8), a concordance program developed by Anthony (2019) to assist in identifying and calculating the lemmas (word families) that are present in the analyzed texts.

In analyzing the data, the texts were marked and tabulated for their lexical items (content words), grammatical items (function words), ranking clauses, and clause complexes. In a clause, the tense was regarded as representing one verb. Moreover, a clause consists of one predictor for the ratio in determining the grammatical intricacy. Only ranking clauses are included in the lexical density analysis, the embedded clauses are not counted. The reason why embedded clauses are not included in the measurement is that embedded clauses are at a lower rank than clauses. Ravelli (1999) states that the relationships formed in embedded clauses are between parts of the clause, not between whole clauses. Thus, embedded clauses

are excluded from the analysis of density and intricacy (Ravelli, 1999).

III. FINDINGS AND DISCUSSION

This study has been examined the text complexity through three different measurements, they are lexical density, lexical variation, and grammatical intricacy. Identification of elements that build the lexical density, lexical variation, and grammatical intricacy was conducted qualitatively based on the Systemic-Functional Linguistic framework. The findings are presented as follows.

Lexical density helps to identify the complexity of text by measuring the ratio of lexical items or content words (nouns, verbs, adjectives, and adverbs) to ranking clauses. According to Halliday (2014), lexical density can be measured by dividing the number of lexical items by the number of ranking clauses. The number of lexical items per clause should be considered for the calculation rather than the total number of words. Therefore, in this study, the ratio of the lexical density was calculated by dividing the number of lexical items with the number of ranking clauses in the text.

The lexical items, the functional items, and the ranking clauses were measured based on the formulas presented earlier. The calculation results are presented in Table I and Figure 1..

TABLE I. LEXICAL DENSITY FEATURES WITHIN TEXTBOOKS

Features	Textbooks					
	VII		VIII		IX	
Semester	1	2	3	4	5	6
Words	40	107	81	45	57	587
Functional items	24	140	42	26	14	298
Lexical items	15	46	39	19	45	235
Ranking clauses	8	19	16	6	13	100
Lexical density	1.87	2.42	2.44	3.17	3.46	2.35

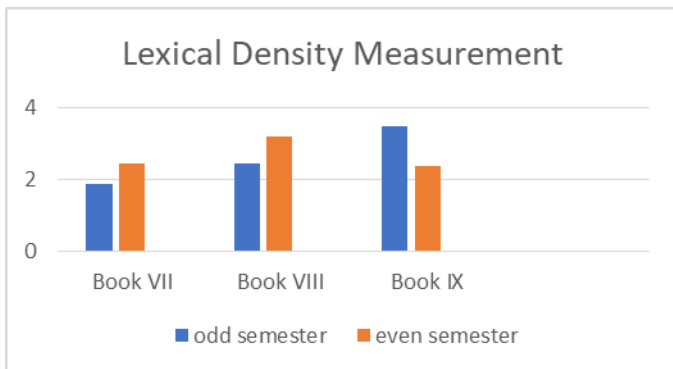


Fig. 1. Lexical Density Measurement

It can be seen that lexical density demonstrated up and down results from the low to the high level of textbooks. Specifically, lexical density in the book VII is developed from 1.87 in the first semester to 2.42 in the second semester and displayed an average number of 2.18 lexical items per ranking clause. In the book VIII, the lexical density of texts is increased from 2.44 in the 3rd semester to 3.17 in the 4th semester and displayed an average number of 2.5 lexical items per ranking

clause. This score rose gradually from 2.18 in book VII to 2.5 in book VIII before slightly dropped to 2.48 in the upper grade that is in book IX. The lexical density in the first semester (semester 5) of this book is higher compared to the lexical density in book VIII which indicate the development of lexical density from book VIII to book IX, but in the last semester, which is semester 6, the lexical density decreases from 3.46 in semester 5 to 2.35 in semester 6.

TABLE II. LEXICAL DENSITY FEATURES ACROSS TEXTBOOKS

Features	Textbooks		
	VII	VIII	IX
Words	147	126	644
Functional items	164	71	312
Lexical items	59	55	280
Ranking clauses	27	22	113
Lexical density	2.18	2.5	2.48

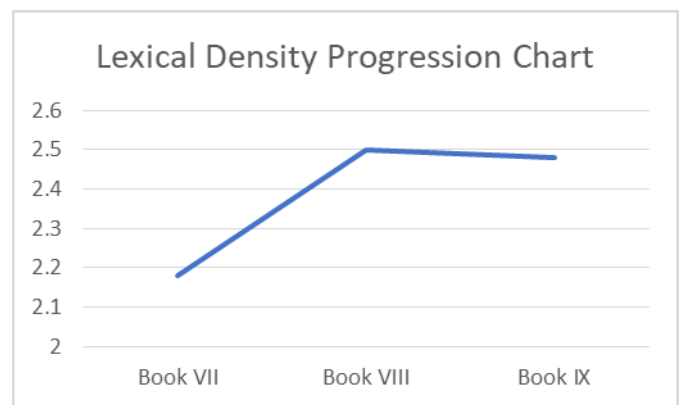


Fig. 2. Lexical Density Progression Chart

Table II and Figure 2 shows a massive increase in the lexical density index from grade VII to grade VIII, while from grade VIII to grade IX, the level of lexical density decreases. It shows a relatively low level of text complexity in grade IX compared to grade VIII. However, even though in grade IX the lexical density decreases, the decrease is not significant. Book IX had a chance to be perceived as more difficult if the lexical items in book VIII are repeated several times, while in book IX, such items are mentioned once (Putra & Lukmana, 2017). The results will be known in the lexical variation measurement below (Table III & Figure 3).

TABLE III. LEXICAL VARIATION FEATURES WITHIN TEXTBOOKS

Features	Textbooks					
	VII		VIII		IX	
Semester	1	2	3	4	5	6
Words	40	107	81	45	57	587
Functional items	24	140	42	26	14	298
Lexical items	15	46	39	19	45	235
Lemmas	14	34	33	18	35	151
Lexical variation	0.93	0.74	0.85	0.95	0.78	0.64

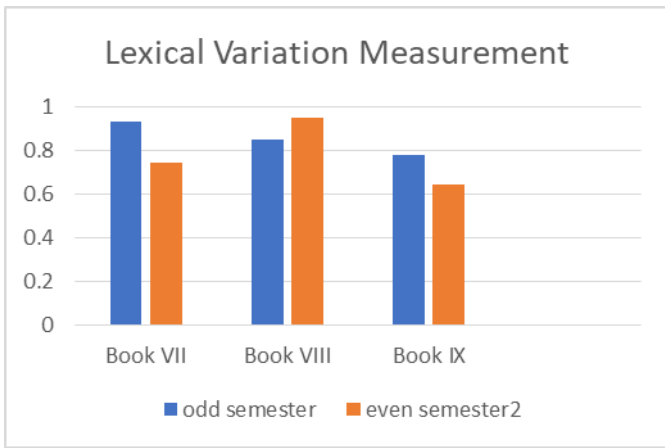


Fig. 3. Lexical Variation Measurement

Lexical variation is another important measurement of text complexity. Halliday (1989) states that the repetition of words creates a sense of reduction that affects the complexity of the text. It means that if there is an approximately equal number of grammatical and lexical items in two texts, but in one text some lexical items are repeated several times, while in the other such items are used only one time, then the latter is likely to be perceived as more complex or difficult. In this study, the number of words in text 6 is much higher than the number of words in other texts. It seems that text 6 has a lower lexical density to make it (the longer text) considered to have a level of difficulty that is not much different from the short texts contained in the three textbooks.

TABLE IV. LEXICAL VARIATION FEATURES ACROSS TEXTBOOKS

Features	Textbooks		
	1	2	3
Functional items	152	71	312
Lexical items	58	55	280
Lemmas (word families)	48	51	186
Lexical variation	0.83	0.93	0.66

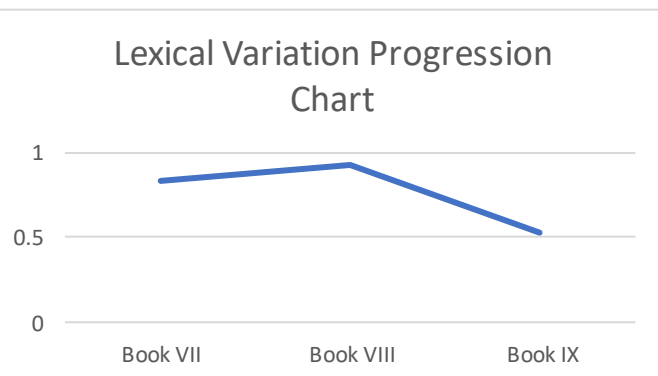


Fig. 4. Lexical Variation Progression Chart

From the data shown in Table IV., book VII has 0.83 of lexical variation and it increases up to 0.93 in book VIII. Thus, the message contained in the former book is likely to be perceived easier to understand than that contained in the latter

one. More variation of the lexical items in a text indicates that there is more information provided in a text which can be grasped by the readers. Moreover, book VIII has 51 lemmas (word families) and 55 lexical items. While textbook IX, has 186 lemmas, 280 lexical items, and 0.66 lexical variations. To be clearer, Figure 3. and Figure 4. show the level of lexical variation from one textbook to another. Figure 3. shows an ups-and-downs progression in the lexical variation index from the first semester until the last. A different trend can be noted concerning book IX, which has a lexical variation index 0.66. This index is 0.27 lower than the index of lexical variation of book VIII and 0.17 lower than book VII. A high lexical variation indicates a highly varied use of lexical items in one clause. In contrast, a low lexical variation indicates a relatively low variation in the use of lexical items in a clause. The level of lexical variation decreases in book IX.

These results indicate that the texts in the textbooks provided by the government still needs improvement since there are so many “ups-and-downs” scores in the lexical variation measurement. Therefore, these fluctuating results indicate the complexity of the texts does not correspond with the level of students’ academic development, or in other words, the direction of the progression does not in line with the grades or levels of the textbooks. This phenomenon, in general, is in line with that of lexical density results, and both of them do not correspond with what has been proposed in Bloom’s taxonomy that has become the worldwide academic expectation. The academic expectation is that students grasp new knowledge through language as their level of education progresses (Bloom et al., 1956, as cited in Schleppegrell, 2004).

Different from the lexical density and lexical variation measurement, the grammatical intricacy to some extent lessens the effect of the lexical complexity (Fang, 2005). It means that the bigger the score in grammatical intricacy, the lesser the complexity of the text is.

TABLE V. GRAMMATICAL INTRICACY WITHIN TEXTBOOKS

Features	Textbooks					
	VII		VIII		IV	
Semester	1	2	3	4	5	6
Clauses	8	19	16	6	13	100
Clause simplexes	8	13	14	2	6	12
Clause complexes (CC)	0	3	2	2	3	33
Ranking clauses in CC	0	6	4	4	8	88
Grammatical intricacy	0	2	2	2	2.67	2.67

TABLE VI. GRAMMATICAL INTRICACY ACROSS TEXTBOOKS

Features	Textbooks					
	VII		VIII		IV	
Semester	1	2	3	4	5	6
Clauses	8	19	16	6	13	100
Clause simplexes	8	13	14	2	6	12
Clause complexes (CC)	0	3	2	2	3	33
Ranking clauses in CC	0	6	4	4	8	88
Grammatical intricacy	0	2	2	2	2.67	2.67

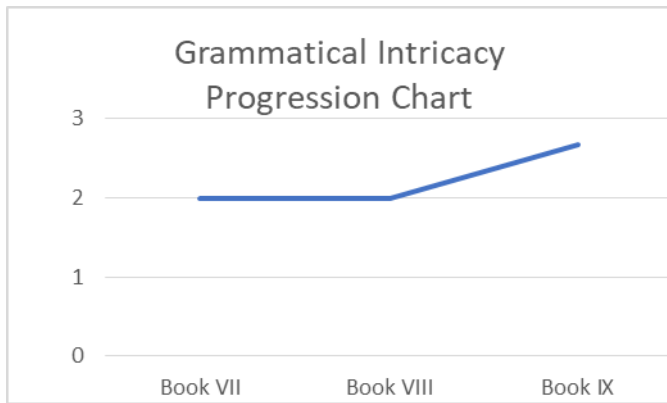


Fig. 5. Grammatical Intricacy Progression Chart

Since the progression of grammatical intricacy indicates a decrease in text complexity, the results shown in Table V, Table VI, and Figure 5 tell that the complexity of the texts decreases from grade VIII to grade IX, and show the same scores between book VII and book VIII. It seems that the decrease in the complexity of the text was intentional so that the longer text in class IX was at the same level of difficulty as the short texts previously taught. Text 6 is the only long text in this series of books. The difference is so significant that it requires a lot of repetition and also a minimized number of lexical items. This makes sense because the length of the text alone makes it difficult for students to read the text. Therefore, long texts are made simple to balance the level of complexity with the short texts contained in the books.

#### IV. CONCLUSION

The findings show that the index of lexical density among the three textbooks increased from the first book to the second book, then decreased in the third book. In other words, lexical density is not in line with the level of students' academic development. However, this conclusion can still be disputed because the level of reduction in lexical density in the third book is far less than the very significant increase in lexical density in the second book.

From the findings above, there are two possible scenarios, first, that the textbook writers are not aware of the effect of lexical density, lexical variation, and grammatical intricacy on a text since the results of the two analyses show the fluctuating results and latter one shows the decreased results. Another possible scenario is that the writers deliberately promote abstraction through lexical density and lexical variation in the shorter texts and the longer text lessens it through grammatical intricacy and also through lowering the lexical density and the lexical variation. In other words, the writers deliberately made long text simple to balance the level of complexity with the short texts contained in the textbooks.

In conclusion, in contrast with Putra and Lukmana's (2017) findings, the texts in the junior high school textbooks published by the Indonesian Ministry of Education and Culture need to be reviewed. The authors must be able to adjust the proportion of

the length of the texts so that the number of words between the texts does not differ significantly. That way, the text can be measured fairly and can be measured scientifically by taking into account the number of lexical items, the number of repetitions of lexical items, the number of ranking clauses, and clause complexes in a text.

#### ACKNOWLEDGMENT

Universitas Perjuangan Tasikmalaya, particularly Department of English Education, is gratefully acknowledged for supporting the research.

#### REFERENCES

Anthony, L. (2019). AntConc Windows (3.5.8) [Computer Software]. Available from <https://www.laurenceanthony.net/>

Apple, M. (1991). The culture and commerce of the textbook. In M. Apple and L. Christian-Smith (Eds.), *The politics of the textbook* (pp. 147-162). New York: Routledge.

Castello, E. (2008). *Text complexity and reading comprehension tests*. Bern, Switzerland: Peter Lang.

Dole, J.A. & Osborn, J. (2003). Elementary language arts textbooks: A decade of change. In J. Flood, D. Lapp & J. R. Squire (Eds.), *Handbook for research on teaching the English Language Arts* (pp. 631-639). New York: Macmillan.

Elliot, D. & Woodward, A. (Eds.). (1990). *Textbooks and schooling in the United States* (90th yearbook of the society for the study of education). Chicago: University of Chicago Press.

Fang, Z. (2005). Scientific literacy: A systemic functional linguistics perspective. *Science education*, 89(2), 335-347. doi: 10.1002/sce.20050

Freebody, P. (2003). *Qualitative research in education: Interaction and practice*. London: Sage.

Halliday, M. A. K. (1985a). *An introduction to functional grammar*. London, England: Edward Arnold.

Halliday, M. A. K. (1985b). *Spoken and written language*. Geelong, Victoria, Australia: Deakin University.

Halliday, M.A.K. (1989). *Spoken and written language*. Oxford: Oxford University Press.

Halliday, M.A.K. (1987). Spoken and written modes of meanings. In R. Horowitz and S. J. Samuels (Eds.), *Comprehending Oral and Written Language* (pp. 55-82). San Diego: Academic Press.

Halliday, M. A. K. (2008). *Complementarities in language*. Beijing, China: The Commercial Press.

Halliday, M.A.K. & Martin, J.R. (1993). *Writing science: Literacy and discursive power*. London: University of Pittsburgh Press.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *An introduction to functional grammar* (4th ed.). London, England: Routledge.

Kwapien, J.J., Drozd, S.S. & Orczyk, A.A. (2010). Linguistic complexity: English vs. Polish, text vs. corpus. *Acta Physica Polonica, A.*, 117(4), 716-720. doi: 10.12693/aphyspola.117.716

Martin, J. R. (2002). Meaning beyond the clause: SFL perspectives. *Annual Review of Applied Linguistics*, 22, 52-74. doi: 10.1017/s026719050200003x

Mukundan, J., Hajimohammadi, R., & Nimehchisalem, V. (2011). *Developing an English language textbook evaluation checklist*. Contemporary Issues in Education Research. 4(6), 21-27 doi: 10.19030/cier.v4i6.4383

Nathan, M.J., Long, S.D. & Alibali, M.W. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of textbooks. *Discourse Processes*, 33(1), 1-21. doi: 10.1207/s15326950dp3301\_01

- Newnham, S. (2013). *Text complexity in graded readers: A systemic functional look* (Master's thesis). Retrieved from <http://mds.marshall.edu/etd/498/>
- Ravelli, L. J. (1999). *Metaphor, mode and complexity: An exploration of co-varying patterns*. Nottingham, UK: Nottingham Trent University.
- Palinscar, A.S. & Duke, N.C. (2004). The role of text and text/reader interactions in young children's reading development and achievement. *Elementary School Journal*, 105(2), 183-198. doi: 10.1086/428864
- 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. doi: 10.17509/ijal.v7i2.8352
- Reys, B.J., Reys, R.E. & Chávez, O. (2004). Why mathematics textbooks matter. *Educational Leadership*, 61(5), 61-66. doi: 10.1111/j.1949-8594.1999.tb17449.x
- Rimmer, W. (2008). Putting grammatical complexity in context. *Literacy*, 42(1), 29-35. doi: 10.1111/j.1467-9345.2008.00478.x.
- Schleppegrell, M.J. (2004). *The Language of schooling: A functional linguistic perspective*. Mahwah: Lawrence Erlbaum Associates.
- To, V. (2017). Grammatical intricacy in EFL textbooks. *The International Journal of English Language Education*, 5, 127-140. doi: 10.5296/ijele.v5i2.12087
- To, V., Fan, S., & Lê, Q. (2016). Research writing: A systemic functional linguistics perspective. In N. Fitzallen, S. Fan, & R. Reaburn (Eds.), *What's next in Educational Research* (pp. 341-352). Rotterdam, The Netherlands: Sense Publishers.
- To, V. (2018). *Linguistic complexity analysis: a case study of commonly-used textbooks in Vietnam*. SAGE Open July-September 2018: 1 –13 doi: 10.1177/2158244018787586
- Watts-Taffe, S., Gwinn, C.B., Johnson, J.B. & Horn, M.L. (2003). Preparing pre-service teachers to integrate technology with the elementary literacy program. *The Reading Teacher*, 57(2), 130–138. doi: 10.1163/9789087907532\_009