



The Readability of Malaysian English Children Books: A Multilevel Analysis

Adlina Ismail (Corresponding author)

Faculty of Languages and Communications, Universiti Sultan Zainal Abidin, Gong Badak Campus, Malaysia

E-mail: adlina1989@gmail.com

Normah Yusof

Faculty of Languages and Communications, Universiti Sultan Zainal Abidin, Gong Badak Campus, Malaysia

Kamariah Yunus

Faculty of Languages and Communications, Universiti Sultan Zainal Abidin, Gong Badak Campus, Malaysia

Received: 03-06-2016

Accepted: 09-08-2016

Advance Access Published: September 2016

Published: 01-11-2016

doi:10.7575/aiac.ijalel.v.5n.6p.214

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.5n.6p.214>

Abstract

These days, there are more English books for children published by local publishers in Malaysia. It is a positive development because the books will be more accessible to the children. However, the books have never been studied and evaluated in depth yet. One important factor in assessing reading materials is readability. Readability determines whether a text is easy or difficult to understand and a balanced mix of both can promote learning and language development. Various researchers mentioned a multilevel framework of discourse that any language assessment on a text should take into account. The levels that were proposed were word, syntax, textbase, situation model and genre and rhetorical structures. Traditional readability measures such as Flesh Reading Ease Formula, Gunning Readability Index, Fog Count, and Fry Grade Level are not able to address the multilevel because they are based on shallow variables. In contrast, Coh-metrix TERA provided five indices that are correlated to grade level and aligned to the multilevel framework. This study analyzed ten Malaysian English chapter books for children using this Coh-metrix TERA. The result revealed that the Malaysian English children books were easy in shallow level but there was a possible difficulty in textbase and situation model level because of the lack of cohesion. In conclusion, more attention should be given on deeper level of text rather than just word and syntax level.

Keywords: Readability, assessment of reading materials, children books, Coh-Metrix

1. Introduction

Readability is defined as the level of ease or difficulty which a text can be comprehended by the readers. Readability is generally important in communicating written information because communication will be hampered if the readability of the text is beyond the readers' comprehension. Readability has also become a major concern for language teachers to choose suitable reading materials for the students. Suitable reading materials would promote the readers' language developments, while too easy text will bore the readers and too difficult text can bring down the readers' motivation (Carrell, 1987; Graesser, McNamara, Cai, Conley, & Pennebaker., 2014).

Because of this, there are a lot of readability formulas introduced in attempts to provide reliable measurements of readability. However, most of them are only mathematical calculations based on word difficulty and sentence length. While they appear to be reasonable and easy to use, these formulas are criticized for their incomplete representation of readability (McNamara Graesser, McCarthy, & Cai, 2014). Even though words choice and sentence structures are crucial in deciding readability, they can only address surface levels of discourse. Deeper levels of discourse which also help in meaning construction can only be represented by other text factors such as noun cohesion, and genre. Coh-Metrix TERA is an automatic tool that is designed to measure deeper characteristic of discourse alongside the shallow variables (word and sentence level). Coh-Metrix Common Core Text Ease and Readability Assessor (TERA) is able to represent readability according to the multilevel theory of discourse.

1.1 Statement of problem

Narrative texts such as story books are especially popular among young readers because children have a tendency to like something with creative and imaginative element (Hamid & Hassan, 2012). Moreover, story books were proven to offer a far wider range of vocabulary than textbooks (Hsieh et al., 2011). Due to these reasons story books are very popular among young readers and have become important tools in language and cognitive developments. In Malaysia, English reading materials are produced locally to reduce over-reliance on reading materials sourced from foreign countries and to implement our own nation's values into children and help them to learn English in Malaysian context

rather than foreign context. These days, there are more locally published children story books in the market compared to a decade ago.

However, the quality of Malaysian children book is questioned. Other than very limited studies were conducted on these reading materials, a few studies had address dissatisfaction with Malaysian children books. Shamsuddin (2003) found that the problem in Malaysian children book publication was due to the lack of quality and they were unsuitable for the target audience, which are children. Moreover, Chew and Ishak (2010) supported this by claiming that Malaysian children literature are not appealing, lacking in quality, is not systematic and not classified according to the children's age. They are also not even in line with the current national development. Reading Promotion Policy was initiated by the Malaysian government to encourage reading habits among Malaysians. However, a part of the policy's objective which is to build English language reading culture among primary schools students is yet to be achieved and one of the contributing factors to the situation is the unsuitable reading materials used in the class (Latiff Azmi, 2013). Furthermore, a recent study on children picture books in Malaysia discovered that the picture books were not highly cohesive which may disrupt readers' deeper understandings (Ismail & Yusof, 2016). Therefore, there is a need for deeper evaluation of Malaysian English reading materials.

1.2 Purpose of the Study

This study aims to analyze the readability of Malaysian English children books according to the multilevel theory of discourse. The five level of discourse that will be discussed in this study are word, syntax, textbase, situation model, and genre (Graesser & McNamara, 2011). This can be achieved with the help from Coh-Metrix TERA.

2. Literature Review

2.1 Children books

Children books basically refer to books which are written for children reading purposes (Mohamed Isa, 2009). There are many types of children books such as board books, picture books, picture story books, rebus books, chapter books, novels and short stories collections. Even then, there is yet any standard method to categorize these books. It is to be remembered that the types of children book sometimes overlap each other. Among the children books, story books had been chosen to be studied. It is because according to Hill (2008), "only books provide the quantity of text in a form that can be read comfortably, and only fiction provides the type of text that can develop a learner's fluency" (p.186). Story books or narrative texts are known to be easier compared to informational texts because the language use is more similar to oral or spoken language. One advantage of story books is that it lets the readers experience events or emotions through the story but if it gets too intense, readers are free to pull out (from the experience or event in the story) and only participate as observers (Ayob, 1981). This will allow the readers to engage in acceptable experiences and events only, unlike in real life. Historically, other than becomes a source of emotional and cognitive enrichment (Chew & Ishak, 2010), story books has been used to teach children about the diversity of society, religion diversity, cultural values and history (Malu, 2013).

2.2 Traditional Readability Measures

Since the 1920s, there were many attempts on predicting text difficulty using shallow indices such as vocabulary and sentence length have been published. Nonetheless, progress and research on the formulas were popularized in the 1950s by researchers like Rudolf Flesh, George Klare, Edgar Dale and Jeanne Chall. Formulas such as Flesh Reading Ease Formula, Gunning Readability Index, Fog Count, and Fry Grade Level were introduced (DuBay, 2004). These formulas were widely used as readability measurements. However, these traditional readability formulas had received a lot of criticizing because of their shallow assumption which analyzes limited features of text only (McCarthy, Lewis, Dufty, & McNamara, 2006; McNamara et al., 2014).

McNamara et al., (2014) pointed out three reasons that traditional readability are not very useful both in terms of theoretically and practical. First, uni-dimensional representations of comprehension tend to ignore the importance of readers' deeper levels of understanding. Shallow indices and cloze tasks that become the foundation of the formulas are not able to explain the deeper level of discourse than word and sentence level. Second, uni-dimensional measures overlook the multiple factors that influence comprehension that influence readers' use of knowledge and deep comprehension such as cohesion and text genre. Third, uni-dimensional measures of text difficulty are not helpful or informative to educators in identifying the difficulty in the text. They only describe the difficulty level of a text, but there is information given in what makes the text difficult or easy. Moreover, depending on shallow variables such as word length and sentence length only may create short, choppy sentences that minimize cohesion and makes the comprehension of the text more difficult (McNamara, Louwerse, & Graesser, 2002).

2.3 The Multilevel Theory of Discourse and Coh-Metrix TERA

The rise of cognitive models of reading has put a spotlight on the need for a measure that accounts for discourse-specific factors such as textbase and situation level processing rather than just surface level (word level). These multiple levels of comprehension have been identified and explored by numerous discourse researchers over the years. Graesser and McNamara (2011) specifically presented a multilevel theoretical framework based on studies of reading comprehension in various fields which includes six levels of discourse (text and spoken), where comprehension can happen. They are words, syntax, explicit textbase, referential situation model, genre and rhetorical structure, and pragmatic communication. The last level will not be included because it is not within the scope of the study. The framework addresses both lower-level of comprehension (basic reading components) such as phonology, morphology,

and word decoding and also higher-level (deeper comprehension components) such as sentence interpretation, construction of inferences, use of background knowledge, reasoning, and knowledge of discourse structures (McNamara et al. 2014).

Formulas which only based on shallow variables are not able to represent the multidimensional framework. However, with the development in technology and cognitive science, Coh-Metrix TERA made it possible to analyze the multilevel of discourse computationally. Coh-Metrix TERA is a computational text assessing tool which is developed based on theory psycholinguistic and cognitive models of reading (Crossley, Allen & McNamara, 2011). It provides more than 60 indices in various levels such as words per sentence, logical connectors, modifiers per noun phrase, word concreteness, and cohesion. According to the previous study (McNamara Graesser, & Kulikowich, 2011), it was found out that five indices (narrativity, word concreteness, syntactic simplicity, referential cohesion and deep cohesion) are corresponded to the grade levels and at the same time are able to represent the five levels of discourse proposed by Graesser and McNamara (2011). These indices were incorporated into a new version of Coh-Metrix, which is Coh-Metrix Text Ease and Readability Assessor (TERA). Coh-Metrix TERA analyses text according to the five variables separately. The component score is presented in percentage figure and sometimes referred as Ease Score. The term ease is used because the five indices (narrativity, word concreteness, syntactic simplicity, referential cohesion and deep cohesion) illustrate components that help reading comprehension. Next, the five levels of discourse will be discussed in related to the indices of Coh-Metrix TERA.

2.3.1 Words

At word level, it is believed that each word in the text is decoded by the reader separately. In Coh-Metrix TERA, Word Concreteness index represents this surface level. As concrete concepts refer to perceivable and spatially embedded entities, abstract concepts refer to entities that are neither purely physical nor spatially constrained (Caramelli, et al., 2004). Texts with concrete words are found to be more comprehensible and interesting than abstract texts (Shilfhout, 2014). In Dual Code theory by Paivio (1986), it explains that concrete words are accessed (by the brain) more easily than abstract words because the information they convey rests on both a verbal and an imagery code, while that conveyed by abstract words rests only on the verbal code. Because of that, concrete words are usually acquired earlier than abstract words by both second language learners and native learners (Rafi, 2013; Caramelli et al., 2004). Concreteness influences children's mental lexicon as that they can comprehend concrete words faster than abstract (Rafi, 2013). Therefore, texts that contain more abstract words are more challenging to understand for young readers.

2.3.2 Syntax

After the word is recognized (by meaning), syntactic parsing happens to link the words together into meaningful phrases and constituents (Crossley & McNamara, 2011). At this level, comprehension is influenced by syntax structure. Many traditional readability formulas such as Flesch-Kincaid Grade Level, Gunning Fog Index and Fry Graph grade texts are based on syntactic complexity. It is because complex sentence structure place higher demands on memory processing, especially for less skilled readers. Hence, they will need to spend more time to understand the sentence (McNamara et al., 2014). While sentence complexity can represent text difficulty, determining syntax complexity only based on word count per sentence or sentence length which decided by traditional formulas is not valid (McNamara, Crossley & McCarthy, 2010). On the other hand, Syntactic Simplicity index in Coh-Metrix TERA analyses sentence complexity by word count, clause count, sentence count, and sentence structure. Other than containing simple sentence structure, text with similar sentence structures also is considered easy in syntactic simplicity. Similar syntactic structure can result in lower cognitive demands on the reader and more attention can be paid to meaning (Crossley et al., 2011). Second language readers and less skilled readers might be affected by syntactic complexity more than first language and skilled readers because they are more unfamiliar to a various syntactic structure of the language (Crossley & McNamara, 2011).

2.3.3 Textbase

Next, at textbase level, connections need to be made between these discourse constituents at the sentence level, paragraph level, and text level in order to develop a coherent mental representation of the text (Crossley & McNamara, 2011). These connections can be enhanced by the presence of cohesion devices that provide links among textual elements and develop coherence for readers such as referential cohesion. Referential cohesion is the overlaps of nouns, pronouns or noun-phrase arguments in the text. They help establishing coherence at textbase level by connecting the content words in a sentence to words in surrounding text or sentences (McNamara et al, 2014). Coh-Metrix analyzes referential cohesion by computing major types of lexical coreference types; noun overlap, pronoun overlap, argument overlap, stem overlap and content word overlap. This overlap in a text can reduce the need to make an inference as the sentences will be very clear. Cohesion devices in a text are especially important for low-knowledge readers such as children because they are less capable of making inferences compared to high-knowledge readers (Schwanenflugel & Knapp, 2015).

2.3.4 Situation Model

The situation model is the subject matter that is being described in informational text or the "microworld" that evolves in a narrative text (McNamara et al., 2014). Situation model refers to the deeper meaning representations that involve much more than the explicit words (McNamara & Graesser, 2012). Any breaks in coherence will require cohesion devices such as connectives words, adverbs, transitional phrases or other signaling devices to stitch them up together again. These cohesion devices help comprehension by revealing to the readers that there is a discontinuity in the text. This level is measured by Deep Cohesion Index in Coh-Metrix TERA.

Sentences in beginner texts are usually shortened to reduce syntax complexity. However, it is usually done by crossing out coherence markers such as connectives and signaling phrases to reduce the word count. Yet, it does not mean the text is easier to read as it can result in coherence gaps and readers will have to determine the type of coherence relation between the clauses by themselves (Shilfhout, 2014). Even though sometimes low cohesion is desirable to encourage inference constructing which will lead to better comprehension (McNamara et al., 2014), children are weak in making inferences. This is due to two reasons; they often lack the prior knowledge upon which such inferences would be based on and they may not be able to retrieve relevant information from prior knowledge even when they possess it (Schwanenflugel & Knapp, 2015). So, the absence of these signaling devices can be a problem to children who read independently.

2.3.5 Genre

Narrative text is known to be easier than expository text even though there is no clear-cut categorization of the genre (McNamara, Graesser, & Louwerse, 2012; Shilfhout, 2014). This claim is supported by McNamara et al. (2014) which stated that narrative text is usually easier for inexperienced readers or readers with low world-knowledge. This is due to the reason that comprehension is enhanced by increased knowledge and narrative text is more within children's knowledge (or familiar) compared to science text. The narrative text tells a story, with characters, event, places, and things that are familiar to the reader. It is very similar to a spoken conversation and closely affiliated with word familiarity, world knowledge, and oral language indices (Graesser et al., 2014; McNamara & Graesser, 2012). Narrativity index analyses the extent to which a text is classified as oppose to informational genre or the degree to which the text conveys a story, a procedure, or a sequence of episodes of actions and events with animate beings (Graesser, McNamara, & Kulikowich, 2011). McNamara et al. (2012) provided an empirical study that shows narrative text is constructed of easy words (familiar and content), complex sentence structure, high in both referential and verb cohesion and not many connectives. The rhetorical structure is the organization of the text at a macro-level and the discourse function of particular excerpts (McNamara et al., 2014). The rhetorical structure is closely related to text types and by understanding the rhetorical structure, readers will be able to comprehend the text better.

3. Methodology

3.1 Samples

Children books that were chosen for this study were children chapter books. Compared to picture books, which are also popular among children, text in chapter books are not dependent on the pictures in the books. They are there solely to appeal to the children. Hence, studying the text separately from the picture will not affect the meaning construct of the story. Chapter books are suitable for older children, nine years old and above (Mohamed Isa, 2009) as they are able to stop and continue the story in some other times. For this study, ten children chapter books were collected from public libraries. Each chosen chapter book contains only one story but is divided into a few short chapters. The books have to be published by Malaysian publishers in order to be selected.

3.2 Procedure of Data Analysis

After the samples were collected, Optical Character Recognition (OCR) software was used to help recognizing words and letters from picture (.jpg, .tnt, and etc.) format and turns them into text (.txt) format. In this process, only narrative texts will be taken as data. Other words and numbers which are not a part of the narration such as contents guide, chapter title, page numbers, glossary, comprehension questions and notes will not be included because they will affect Coh-Metric T.E.R.A indices scores. Manual editing was done on the accordingly without changing the original text.

Next, the texts will be evaluated in online Coh-Matrix T.E.R.A (<http://129.219.222.66:8084/Coh-Matrix.aspx>) for five variables; narrativity, syntactic simplicity, word concreteness, referential cohesion and deep cohesion (McNamara, Graesser, Cai, & Dai, 2013). All of the chapters in a book are inserted as a long text, divided into paragraphs. Coh-Matrix T.E.R.A score can also be called as *easability* score because as oppose to difficulty, it measures how easy the text is to be comprehended. The higher percentage of the result shows the text is easier to be comprehended. The result will be in percentage for each variable and will be represented in bar charts with a narrative description.

4. Analysis and Discussion

The tool, Coh-Matrix TERA describes percentage below 25 as low, above 75 as high and in between as average. The range for the average score has the biggest range compared to the high and low score which only cover the 25% of top and bottom percentage respectively. The result of the Ease Score for all ten samples is summarized in a frequency distribution table below.

Table 1. Frequency distribution

	Narrativity	Syntactic Simplicity	Word Concreteness	Referential Cohesion	Deep Cohesion
High (85-100%)	9	10	2	0	1
Average (26- 74%)	1	0	7	5	9
Low (0- 25%)	0	0	1	5	0
Total of sample size	10	10	10	10	10

There was no missing data. The data from all ten samples were included in the study. According to the frequency distribution, the majority of the sample (90%) scored high in Narrativity. This means that the language use is very close

to oral communication with a lot of high-frequency words. This was expected because the samples chosen are all story books. For the syntactic simplicity, all (100%) samples present high score in this index. The samples were proven to be dominated by short, easy and similar sentence structures which are easy to be comprehended by beginner readers. Next, different from the previous indices, most of the samples (70%) were average in word concreteness, implying that the ratio of concrete words to abstract words is moderate. For referential cohesion, the result was divided into two groups, average and low with both have the same frequency ($n = 5$) There were moderate overlaps of nouns, noun phrases and pronouns in half of the sample while the other half had very little of idea overlaps. Lastly, most of the sample scored average in deep cohesion which means that they contained a fair number of connectives, transition signals, and adverbs. To look at the data in more detailed, descriptive statistic in Tab2 is referred.

Table 2. Descriptive analysis of Coh-Metrix TERA indices

	Minimum	Maximum	Mean	Std. Deviation
Narrativity	70	96	85.30	8.084
Syntactic Simplicity	81	97	88.60	6.501
Word Concreteness	20	83	54.20	20.088
Referential Cohesion	13	53	27.50	12.730
Deep Cohesion	43	80	55.70	11.870

The means of the sample is aligned to the frequency distribution, showing high score in narrativity and syntactic simplicity and average in word concreteness, referential cohesion, and deep cohesion. By looking at the mean of referential cohesion, it is clear that the samples were actually average in referential cohesion but with a borderline score, which was almost low. So, it means the samples are quite low in connectives and transitional phrases content.

Measured as high in ease score illustrates that it is very easy or there is no difficulty at all, while low score implies that it is very difficult to comprehend at that level. This study found that the samples were very easy at the shallow levels (word and syntax) for Malaysian children. In Malaysia, English is taught as the second language. Thus, as second language readers, Malaysian children are still in the process of acquiring the grammar when they start reading English books (Heilman, Collins-Thompson, Callan, & Eskenazi, 2007). Thus, short and easy sentence structures in the texts were not disadvantages, but are potential to aid young readers' comprehensions. The numbers of concrete words were average in the samples but this is seen as suitable for the target readers' age. Mohamed Isa (2009) mentioned that chapter books (in mother tongue) are suitable for children at the age of seven until nine years old, where they also start to understand the abstract concept. For books in second language, they are suitable for slightly older children due to the language barrier. However, the chapter books readers should be at the age where a moderate number of abstract words may not disturb their reading process. In textbase level, there was a significant difficulty presented by the samples. This means there is quite a few of ideas overlaps in the chapter books.

The lack of overlaps in the text can result in cohesion gaps which required inferences to be made in order to connect the ideas between sentences and paragraphs. Referential cohesion cues are particularly important in texts with a less familiar topic (McNamara, Graesse, & Louwse, 2012) and for children who generally have low world knowledge (Schwanenflugel & Knapp, 2015). In situation model level, there was also slight difficulty because of the moderate number of connectives, transition signals, and adverbs. These devices reveal that there are cohesion gaps in the text explicitly and at the same time provide information of the relationship between the sentences or ideas. In narrativity level, there was no difficulty because the samples shared a lot of narrative features which are very friendly to young second language readers. By comparing to the previous research on Malaysian English picture story books (Ismail & Yusof, 2016), the readability of chapter books is almost similar to picture books. They were only different in terms of the density of concrete words. There were more abstract words in chapter books compared to picture books. Although both of types of children books contained a lot of familiar words and simple sentences, both of them were not highly cohesive. While cohesion can be built through incorporated illustration in picture books, chapter books construct meaning solely on text.

5. Conclusion

Even though these comprehension components are called levels, they occur simultaneously. It is possible to construct comprehension at the surface level only and not at the deeper level (McNamara et al., 2014). Moreover, higher level processing is not better, or harder; it is just processing that is closer to conscious introspection on the part of the reader (Grabe, 2014). Deeper reading components are more time consuming, strategic, and taxing on cognitive sources of readers.

In conclusion, potential difficulties in Malaysian English children books appeared in deeper levels (textbase and situation model) of discourse only. Readers should be able to gain comprehension at the surface level without any problem because the books contained familiar words, a lot of concrete words and short sentences. In textbase level, the lack of idea overlaps may be due to the writer's confusion between good the quality of reading materials and essay for evaluations (e.g., during examinations). While text with greater lexical diversity in the text is considered more

sophisticated and generally scores higher in writing assessment, it is also more difficult to understand (McNamara et al., 2010). Thus, it is important for writers and editors to consider their targeted readers while developing reading materials for children. Moreover, more attention should be given on cohesion in Malaysian English reading texts because they are especially important to generate comprehension for readers with low domain knowledge (McNamara et al., 2014). This study provides a clearer view of Malaysian English children books was presented. Educators, writers, and publishers should make use of the existing studies on reading materials to improve local products quality. Providing good quality and suitable reading materials can enhance children's language developments.

References

- Ayob, S. (1981). Sastera untuk kanak-kanak: Fungsi dan implikasinya. *Pendidik dan Pendidikan*, 3(1), 73-82.
- Carrell, P.L. (1987). Readability in ESL. *Reading in a Foreign Language*, 4(1), pp.21– 40. <http://nflrc.hawaii.edu/rfl/PastIssues/rfl41carrell.pdf>
- Caramelli, N., Setti, A., & Maurizzi, D.D. (2004). Concrete and abstract concepts in school age children. *Psychology of Language and Communication*, 8(2), 19-34.
- Chew, F. P., & Ishak Z. (2010). Malaysian folk literature in early childhood education. *World Academy of Science, Engineering and Technology*, 66, pp.557-564. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1877050913012209>
- Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 84 -101. Retrieved from <http://nflrc.hawaii.edu/rfl>
- Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life-Long Learning*, 21 (2/3), 170-191. Retrieved from http://www2.gsu.edu/~wwwesl/Scott_Publications.html
- DuBay, W. H. 2004. *The Principles of Readability*. Retrieved from <http://www.impact-information.com/impactinfo/readability02.pdf>.
- Shamsuddin, F. (2003). Masalah penerbitan buku kanak-kanak. *Kertas Kerja Seminar Terbitan Sastera Kanak-Kanak* (pp.22-23). Dewan Bahasa dan Pustaka.
- Grabe, W. (2014). Key Issues in L2 Reading Development. In *CELC Symposium Bridging Research and Pedagogy* (pp.8-18). Retrieved from <http://www.nus.edu.sg/celc/research/symposiumproceedings.php>
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, 3, 371–398. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1756-8765.2010.01081.x/full>
- Graesser, A. C., McNamara, D. S., & Kulikowich, M. (2011). Coh-Metrix: Providing Multilevel analysis of text characteristics. *Educational Researcher*, 40(5), 223 – 234. Retrieved from <http://er.aera.net>
- Graesser, A.C., McNamara, D.S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, 15(2), 210-229.
- Hamid, R., & Hassan, S. (2012). Penerapan sastera dalam pengajaran bahasa di sekolah rendah. In *Seminar Bahasa Melayu 2012: Kecekapan Berbahasa dan Penghayatan Nilai dalam Pendidikan Abad Ke-21* (pp. 226-234). Singapore: MLCS. Retrieved from <https://malaylanguagecentre.moe.edu.sg/images/pdf/16-penerapan-sastera-dalam-pengajaran-bahasa-di-sekolah-rendah-rosnita-sarifah.pdf>
- Heilman, M. J., Collins-Thompson, K., Callan, J., & Eskenazi, M. (2007). Combining lexical and grammatical features to improve readability measures for first and second language texts. In *Proceedings of NAACL HLT* (pp. 460 – 467). Rochester, NY: Association for Computational Linguistics.
- Hill, D. R. (2008). Graded readers in English. *ELT Journal*, 62(2), 184 – 204. Retrieved from <http://eltj.oxfordjournals.org>
- Hsieh, M., Wang, F., & Lee, S. (2011). A corpus-based analysis comparing vocabulary input from storybooks and textbooks. *The International Journal of Foreign Language Teaching*, 25-33.
- Ismail, A., & Yusof, N. (2016). Readability of ESL picture books in Malaysia. *Journal of Nusantara Studies*, 1(1), 60-70.
- Malu, K. F. (2013). Exploring children's picture storybooks with adult and adolescent EFL learners. *English Teaching Forum*, 3, 10-18.
- McCarthy, P. M., Lewis, G. A., Dufty, D. F., & McNamara, D. S. (2006). Analyzing writing styles with Coh-Metrix. In G. C. J. Sutcliffe & R. G. Goebel (Eds.), *Proceedings of the 19th annual Florida artificial intelligence research society international conference* (pp.764-770). Retrieved from <http://aaaipress.org/Papers/FLAIRS/2006/Flairs06-151.pdf>

- McNamara, D., Louwerse, M., & Graesser, A. (2002). Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. Grant proposal. Retrieved from <http://cohmetrix.memphis.edu/cohmetrixpr/publications.html>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27 (1), 57 – 86. doi: 10.1177/0741088309351547. Retrieved from <http://wex.sagepub.com>.
- McNamara, D. S., Graesser, A. C., Cai, Z., & Kulikowich (2011). *Coh-Metrix easability components: Aligning text difficulty with theories of text comprehension*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- McNamara, D. S., & Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P. M. McCarthy & C. Boonthum (Eds.), *Applied Natural Language Processing and Content Analysis: Identification, Investigation, and Resolution* (pp. 188–205). Hershey, PA: IGI Global
- McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring Up: Advances in How We Assess Reading Ability* (pp. 89–116). Plymouth, UK: Rowman & Littlefield Education.
- McNamara, D. S., Graesser, A., Cai, Z., & Dai, J. (2013). Coh-Metrix Common Core T.E.R.A. version 1.0. Retrieved from <http://coh-metrix.commoncoretera.com>.
- McNamara, D.S., Graesser, A.C., McCarthy, P. M. & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. New York, NY: Cambridge University Press.
- Mohamed Isa, Z. (2009). Choosing appropriate Bahasa Melayu books for Malaysian children. *Jurnal Bitara*, 2, 45 – 57.
- Latiff Azmi, M. N. (2013). National language policy and its impact on second language reading culture. *Journal of International Education and Leadership*, 3(1), pp.1-11. <http://www.jielusa.org/wp-content/uploads/2012/01/Azmi-Final.pdf>
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Cambridge University Press. Retrieved from https://books.google.com.my/books?id=hLGmKkh_4K8C
- Rafi, M. S. (2013). Natural order of vocabulary acquisition. *European Academic Research*, 1(5), 721-733. Retrieved from <http://euacademic.org/UploadArticle/52.pdf>
- Schwaneflugel, P. J. & Knapp, N. F. (2015). *The Psychology of Reading: Theory and Applications*. Guilford Publications. Retrieved from <https://books.google.com/books?isbn=1462523501>
- Shilfhout, G. (2014). Fun to read or easy to understand? Establishing effective text features for educational texts on the basis of processing and comprehension research. Retrieved from <dspace.library.uu.nl/bitstream/handle/1874/300805/silfhout.pdf>