

# Developing Readability Computational Formula for Arabic Reading Materials among Non-Native Students in Malaysia

Kamarulzaman Abdul Ghani<sup>1</sup>, Ahmad Sabri Noh<sup>2</sup>, Nik Mohd Rahimi Nik Yusoff<sup>3</sup>  
and Nur Hafezah Hussein<sup>4</sup>

<sup>1</sup> Universiti Malaysia Kelantan, Kelantan, Malaysia

<sup>2</sup> Institute of Teacher Education, Islamic Education Campus, Selangor, Malaysia

<sup>3</sup> Universiti Kebangsaan Malaysia, Selangor, Malaysia

<sup>4</sup> Universiti Malaysia Kelantan, Kelantan, Malaysia  
kamarulzaman@umk.edu.my

**Abstract.** Efficient measures on the readability of Arabic reading materials are important in ensuring quality text to be used. However, there are limited tools in assessing readability of Arabic reading materials. Thus, the lack of measuring tools to assess the suitability of textbooks amongst non-native speakers is a constraint which led to this study. In this regard, this paper discusses the efforts to develop a readability formula for Arabic reading materials in Malaysia. This study involved 313 samples of population from thirteen Arabic textbooks used in religious secondary schools in Malaysia. A linguistic analysis is performed as well as an inference analysis to see the strength of the relationship of each linguistic pattern with the level of readability. The findings show that the relationship between the linguistic features that is being analysed is not strong. But this finding can be used as the basic formula for improvement. This is because the textbooks that have been sampled do not have good writing features. Thus, a comprehensive effort is needed to ensure the preparation of Arabic language textbook is of good quality for non-native students.

**Keywords:** Readability Formula, Readability, Arabic Reading Materials, Religious Secondary Schools, Non-Native, Linguistic Features.

## 1 Introduction

To this day, the practice of reading Arabic text is still being pursued in the discipline of Islamic knowledge as it is a legacy of knowledge passed down from generation to generation through the writings of the *ulama* (Muslim scholars engaged in the field of Islamic studies), a treasure trove to ensure the sustenance of Islamic civilization. The process of transferring knowledge, the formation of thoughts and behaviours are related to the selection of meaningful reading texts. Arabic reading texts serve not only as a medium of language learning, but also as a medium for deepening knowledge based on the indivisible oneness to Allah the Almighty.

Appropriate Arabic reading texts should be selected properly by teachers to gain positive reflection and ensure positive engagement from the students in the teaching and learning. That being said, Arabic readability studies are growing in Arab countries, but still in its early stage in Malaysia. This is important in the Malaysian context due to the limited access and opportunities for using Arabic language. The use of appropriate Arabic reading texts is important in the learning process as it helps to create students' interest and helps them to master the language more effectively [1]. In addition, students need to be trained to have autonomy and responsibility by adopting student-centred learning, such as reading additional reading materials so that they are more independent and active in learning and do not rely solely on teacher support [2].

Consequently, the Arabic reading texts must be at the level of its readability to ensure that the material is compatible with the student's language level.

Measurement on Arabic texts readability and studies relating to Arabic-language readability are still foreign to Malaysia [3], unlike the long-established and developed English-language readability measurement [4]. Therefore, the selection, preparation and usage of Arabic reading materials that suit the level of Arabic language mastery amongst non-native students should be considered to ensure effective continuous reading.

Further, readability measurement has shifted from traditional to modern approach in respond to the recent development in information technology and the widespread use of internet. Readability has been built into search engines and other online tools to facilitate the improvement of overall reading materials' quality [5, 6].

## 2 Studies on the Arabic Texts Readability Measurement

The study [7] investigated the relationship between reading difficulty and specified linguistic variables to determine the effectiveness of the variable in predicting the level of texts readability. 30 excerpts from reading textbooks for grades four, five and six were used. Next, a total of 360 students from grade five answered a cloze test. The test scores were used as criterion variables, while scores on Arabic linguistic variables were used as predictor variables. The linguistic variables studied were average sentence length, average word length, word frequency, percentage of noun phrases and percentage of proper noun. Variables that had significant correlations with the cohort test scores were selected as the predictor factors for the construction of the formula. The study found that the variables with high correlation values of 0.9 were average word length, average sentence length and average word frequency.

Another study [8] was conducted to measure readability. The study aimed to determine the readability of texts using the JAN (Judgment Analysis) method, which is the analysis of expert judgment. A profile score of experts' feedbacks on 60 excerpts selected randomly from Arabic reading textbooks of grade one to grade six primary school students in several Arab countries was used. Fifteen teachers with at least five years of teaching experience in primary schools in Iraq were included as examiners in the study. They were required to rate 60 excerpts based on a ten-point scale; ranging from scale 1 which is easy, suitable for grade one students, to scale 10 which is difficult, suitable for grade six students. Text readability grades are based on a number of linguistic variables. The predictor factors of the text readability included in the stepwise multiple regression equation were:

1. *Average sentence length*
2. *Average word length*
3. *Average word frequency*
4. *Percentage of proper noun*
5. *Percentage of noun phrases*

The findings also showed that there was a significant relationship between the average word length, average sentence length and word frequency. As a result of the multiple regression analysis, the mathematical equation formulas for estimating the readability of a passage are as follows:

$$\text{Policy 1: } RD = 4.41434 (\text{average word length}) - 13.4687$$

$$\text{Policy 2: } RD = 0.98 (\text{average sentence length}) + 0.38 (\text{average word length}) - 0.91 (\text{average number of repeated / frequent words}) - 1.06$$

Empirically, the developed formula can be used manually by teachers to estimate the readability of textbooks provided for the usage of grade 1 to grade 6 students in primary schools in Iraq [8].

Further, a study [9] was carried out by using a simple approach for measuring text readability. The measurement method studied is based on the calculation of

average score of words in the text. Word scores are derived from the frequency of words found in the *King Abdul Aziz City for Science and Technology Arabic Corpus* (KACSTAC), a public corpus containing text from magazines, books, newspapers, journals, dissertations, materials from government agencies, school curriculum and internet materials. The researchers assumed that the more frequent a word is used, the easier it is for the reader to understand the text. Word scores are the inverse values of word rankings in the corpus. Therefore, the lower the value of the overall word score, the easier it is to read the text. However, this study was a preliminary study that did not provide a clear conclusion on the effectiveness of the method.

In addition, an article [10] proposed a method for automatically determining the readability index of Arabic texts by using factor analysis to look at factors that contribute significantly to readability. Mathematical equations from factor analysis are inputted in computer software so that the readability index can be calculated automatically. The researchers used more than 1196 texts from various subjects used by grade one to grade ten students in Jordan. They analysed the linguistic factors of the text: average number of letters in a word, average number of words in sentence, average of difficult words, number of letters, word count, sentence count and difficult word count. However, the factors that contributed to the readability and had highest load were chosen to represent the mathematical equations:

$$\text{AARI Readability Index} = (3.28 \times \text{Number of letters}) + (1.43 \times \text{average word length}) + (1.24 \times \text{average sentence length}).$$

These studies have employed various readability tools or indexes in measuring Arabic texts readability in accordance to their sample and context which is the native Arabic students. Thus, the choice of textbooks and the integration of certain features into their studies was highly based on the students' nativeness. This indicates that there is a lack of readability formulas or tools in the context of non-native students. Thus, there is a need to provide a systematic and scientific approach to the determination of the key features influencing the readability of Arabic texts and include linguistics features that are relevant to non-native students. It is also important to take into account the use of Arabic textbooks and reading materials for non-native students.

### 3 Statement of Problem

The field of readability has long been established and popular in the west [11]. In the context of Arabic reading materials, the attempt to develop a benchmark of its readability is lacking [12]. It is even more difficult to measure the difficulty of an Arabic reading text in the context of non-native readers in Malaysia. The most commonly used measure for benchmarking is by using cloze tests [13]. A review on Arabic-language books used at school and institutions of higher learning in Islamic studies in Malaysia showed that majority of them are written for Arabic native speakers. Even the *al-Azhar* religious curriculum used in religious schools is also Arabic native speakers oriented [1].

This is a critical situation because one of the factors that can help students improve their Arabic reading skills is through the use of reading materials that are appropriate to the student's language level and can increase students' motivation to read them [14]. This can be seen as a weakness for school and university graduates of religious stream or major who are not fluent in Arabic language, particularly in reading.

In addition, there has been no previous attempt to develop a benchmark that can be used to measure the level of difficulty of Arabic reading materials as well as a guide in developing reading materials for Malaysian schools. Thus, a study is needed to establish the basis for developing readability benchmarks for Arabic reading materials in Malaysia [15]. This is in line with the recommendation from [16] in his doctoral thesis which is to conduct a study to identify linguistic variables in the

context of non-native speakers in Malaysia in the development of benchmarks for Arabic readability formula in Islamic studies.

#### 4 Methodology

This study aimed to identify aspects of linguistic variables that influence the degree of difficulty of reading texts among non-native readers in Malaysia, thus establishing a benchmark based on the Arabic text readability formula for non-native readers. The suitable design to achieve the purpose of this study is to use content analysis, which is to analyze the linguistic features used in texts samples. From the data obtained, the study is further developed by using a quantitative approach, in which data is analyzed using descriptive and inferential statistical analysis. This can be done by considering the frequency of a category in the unit of analysis.

The textual content analysis in this study aimed to investigate the textbooks used by Malaysian secondary school students from different levels, ranging from form one to form five. The analysis identified the level of linguistic features' usage in Arabic texts. The linguistic features studied in this study are based on existing studies of readability and have been validated by Arabic linguists. The sample reading texts are 313 texts taken from thirteen Religious Curriculum textbooks used by students from form one to form five in various subjects of Islamic education and Arabic language in Malaysia. The textbooks used are *Mutola'ah* (Prose Literature), *Tafsir* (Commentary and Interpretation of Quran), *Hadith* (Sayings and Teachings of Prophet Muhammad), *Fiqh* (Islamic Jurisprudence), *Tauhid* (Monotheism), *Sirah* (Prophets' History), *Tarikh Islam* (Islamic History), *Adab* (Poetry Literature), *Mantiq* (The science of classical logic), *Tajwid* (Pronunciation), *Balaghah* (Rhetoric), *Nahu* (Grammar) dan *Sorof* (Morphology).

The analysis can provide an accurate representation of the level of usage of Arabic linguistic features. Among the linguistic features studied are simple sentences, complex sentences, noun phrases, common and frequent words, abstract words, conjunctions and punctuation as well as average sentence length. These linguistic features have been identified by researchers in the study of text readability.

Descriptive analysis is performed to determine study variables, frequency, mean and standard deviation [17]. To determine the level of use of linguistic features, this study first determined the rankings of these linguistics features based on the frequency score for each of them. The main inferential statistics used was Pearson's correlation analysis to identify the relationship between linguistic features and readability level as well as multiple regression analysis (stepwise) to ascertain the contribution of linguistic features to text readability level as well as to determine the most significant variable contributing to any changes on dependent variable. The independent variables in this study are the linguistic features that have been identified, while the dependent variables are the text readability level which is in accordance to the students' level. It consists of five levels which are very easy, easy, moderate, difficult and very difficult. The text readability is determined by using a readability scale of 1 to 5, which are: very easy (1) easy (2) moderate (3) difficult (4) very difficult (5).

The definitions and coding procedures were used as guides by the encoder during the encoding process of linguistic features. Table 1 shows a summary of the encoding of the content analysis:

**Table 1.** Summary of linguistic features encoding

Code	Category	Definition	Analysis Procedure
Xp1	Common and frequent words	Common and frequently used words based on students' norms in Malaysia.	Refer to Vocabulary List in [18]. Sum of words in 100 words.
Xp2	Abstract words	Words that convey a general or complex concept, which cannot be proven by the five senses.	Sum of words in 100 words.

Xa1	Simple sentence	The sentence has one independent clause, that is, it can stand on its own to build meaning. Free clauses consist of one subject and one predicate.	Calculate the sum of complete simple sentences.
Xa2	Complex sentence	A combination of one free clause and one or more free clauses. Non-independent clauses cannot stand on their own to build meaning and require explanation.	Calculate the sum of complete complex sentences.
Xa3	Noun phrases	The sentence structure begins with noun.	Calculate the sum of complete noun phrases.
Xa4	Average sentence length	The complete sentence structure builds on the meaning. Sentence consists of subject and predicate or combination of dependent and non-dependent clauses.	Calculate the sum of sentences in 100 words. 100 words are divided by the sum of sentences.
Xh	Conjunction and punctuation	A word or phrase used in a specific function to connect ideas, arguments, stories, concepts with other concepts. It is used to communicate ideas in sentences, between sentences, or between paragraphs.	Sum of conjunction and punctuation in 100 words.

Stepwise multiple regression analysis was used to identify the relative contribution of the independent variable to the dependent variable. Data from each variable are used to view the form of relationships or models that exist, whether linear or non-linear relationships, and to estimate the parameters in the model. The variables used in the analysis to form the relationship can be defined as below:

- Y = dependent variable
- a = is a constant
- b1, b2, b3, bn = regression coefficient
- X1, X2, X3, Xn = independent variables
- r = is the standard error

These variables in the relations can be expressed in a more refined form, namely in the form of multiple linear regression model as follows:

$$Y = a + b1X1 + b2X2 + b3X3 + ..... bnXn + r$$

This equation allows the researchers to obtain a linear combination of independent variables that can be the best predictors of the dependent variable [19].

The validity of the model used is validated by using the adjusted R-squared [20]. The adjusted R-squared value gives an idea of how far the regression model can be generalized to the population. If the value of adjusted R-squared is equal to or close to the R-squared, this indicates that the model is compatible with the data set in the population [21].

In addition to the adjusted R-squared values, the researchers also used the Bootstrap approach in the SPSS software to ensure model validation. The Bootstrap method is an appropriate and more accurate method for testing the validity of the regression model without requiring large sample sizes [22, 23].

## 5 Findings

Descriptive statistics were used to look at the overall use of linguistic features such as the use of simple sentences, complex sentences, noun phrases, common and frequent words, abstract words, conjunction and punctuation as well as the average sentence length. Table 2 shows the overall analysis of these linguistic features.

**Table 2.** Distribution of linguistic features based on overall levels

Variable	n	Mean	S.P
Simple sentence	313	1.99	1.49
Complex sentence	313	5.46	1.25
Noun phrase	313	4.21	2.24
Average sentence length	313	14.18	3.22
Common & frequent words	313	95.44	2.48
Abstract words	313	11.87	5.54
Conjunction & punctuation	313	5.38	2.26

Table 2 shows the mean distribution of the linguistic features studied in the entire texts sample (n = 313). In terms of sentence type, the mean use of complex sentences (mean = 5.46) was 73.24% higher than simple sentence mean value (mean = 1.99) of 26.76%. The mean value of the usage of noun phrases is 4.21 which is 56.56%. Next, the mean value for sentence length is 14.18. In terms of word usage, common and frequent word dominated the texts sample with a mean value of 95.44 which is 95.44% compared to abstract words with a mean value of 11.87 indicating 11.87% of the total 100 words in the text sample. The mean value of use of conjunction and punctuation is 5.38 with 5.38%.

Overall, the results of this data analysis by standard deviation indicate different uniform distribution in the use of linguistic features. The most uniform distribution is in the use of complex sentences followed by simple sentences, noun phrases, conjunction and punctuation as well as common and frequent words, while the distribution for average sentence length indicates less uniform distribution and abstract words indicate non-uniform distribution.

The independent variables in this study were the linguistic features: simple sentences, complex sentences, noun phrases, common and frequent words, abstract words, conjunction and punctuation as well as the average sentence length. The dependent variables consisted text readability levels of 1, 2, 3, 4 and 5. In order to identify the relationships between the variables studied, Pearson correlation tests were conducted to analyse the relationship between these variables.

Table 3 presents the results of the Pearson correlation analysis explaining the relationship between all the variables in the study.

**Table 3.** Pearson correlation between linguistic features and level of text readability

Variable	Level of text readability	
	r	Sig.
Simple sentence	-.102	.072
Complex sentence	-.135	.017
Noun phrase	.027	.637
Common and frequent words	-.606	.000
Abstract words	-.023	.683
Conjunction and punctuation	-.403	.000
Average sentence length	.160	.005

Sig. at the 0.05 level

The strength of the relationship between variables was interpreted based on suggestion in [24] as shown in the table below:

**Table 4.** Interpretation of correlation coefficients [22]

Correlation Value (r)	Correlation level
0.10-0.29	Low
0.30-0.49	Moderate
0.50-1.00	Strong

As can be seen in table 4, variables found to be significantly related to the level of text readability are complex sentence, common and frequent words, conjunction and punctuation as well as average sentence length, while simple sentence, noun phrases and abstract word indicate a non-significant relationship with the level of text readability.

Overall, the results of Pearson's correlation analysis on the variables of linguistic features with the level of text readability indicated that only four variables were significantly correlated, which are, common and frequent words with strong correlations, conjunction and punctuation with moderate correlations, average sentence length and complex sentences with low correlation with the level of text readability. While, no significant relationships were found in the remaining three variables, namely simple sentence, noun phrases and abstract words with the level of text readability.

The statistical procedure (stepwise) is used to obtain a subset of predictor variables (linguistic features) that are useful for predicting criterion variables (text readability levels) in which the order of inclusion of predictor variables into regression equations is based on statistical criteria.

The determinants of predictor variables included in the regression equations and the predictor variables were eliminated based on the correlation values between the predictor variables and the strongest criterion variables. The predictor variables with the lowest correlation values were dropped from being included in the regression equation [20]. The analysis of variance in regression indicates whether the developed model produced a good prediction of its significance to the predictors.

**Table 5.** Analysis of variance

Model		Sum of squares	Degrees of freedom	Mean squares	F	Sig.
3	Residual	279.261	3	93.087	82.959	.000 <sup>d</sup>
	Regression	346.726	309	1.122		
	Sum	625.987	312			

Sig. at the 0.05 level

d. Predictors: (constant), common and frequent words, conjunction and punctuation, average sentence length.

Dependent variable: Level of text readability.

Table 5 shows the ANOVA results showing value of  $F = 82.959$  and its significance value of .000 ( $p < 0.05$ ). Three predictor variables, which are common and frequent words, conjunction and punctuation as well as the average sentence length contribute significantly to the level of text readability.

The results of the stepwise multiple regression analysis that identified the contribution of three independent variables to the text readability level are summarized in table 6. The three variables accounted for 44.6 percent of the variance in the text readability level. Two predictor variables had negative effects of  $-0.298$  (common and frequent words) and  $-0.178$  (conjunction and punctuation). The variable of average sentence length had a positive effect of 0.043.

**Table 6.** Stepwise Multiple regression analysis of independent variables affecting level of text readability

Variable	B	Standard Error	Beta ( $\beta$ )	t	Sig.	R <sup>2</sup>	Contribution (%)
Common and frequent words	-0.298	0.025	-0.522	-11.787	0.000	.367	36.7
Conjunction and punctuation	-0.178	0.027	-0.283	-6.459	0.000	.437	7.0
Average sentence length	0.043	0.019	0.097	2.246	0.025	.446	0.9
Constant	31.830	2.444		13.026	0.000		
R	: 0.668						
R Squared	: 0.446						
Adjusted R Squared	: 0.441						
Constant	: 31.830						
Standard Error	: 2.444						

The most important predictor of linguistic features in the texts sample was common and frequent words ( $\beta = -0.522$ ,  $t = -11.787$  and  $p = 0.000$ ) which accounted for 36.7 percent. This means that as the use of common and frequent words in the text increases, the level of text readability will decrease or in other words the easier the readability level is. This finding also shows that the use of common and frequent words is a major contributing factor to the level of difficulty of text.

The second most important predictor was the conjunction and punctuation that had a negative impact and contributed 7.0 percent to the level of text readability ( $\beta = -0.283$ ,  $t = -6.459$  and  $p = 0.000$ ). This means that as the use of conjunction and punctuation increases in the text, the level of readability decreases. In other words, text that has more conjunction and punctuation indicates easier level of text readability.

The third most important predictor was the average sentence length that had a positive impact and contributed 0.9 percent to the level of text readability ( $\beta = 0.097$ ,  $t = 2.246$  and  $p = 0.025$ ). This means that as the average sentence length in the text increases, the level of readability will also increase. In other words, text that has more average sentence length indicates a high or difficult level of text readability.

In general, the contribution of the three independent variables that are significant to the level of text readability can be made through regression equations. It is a prediction of the dependent variable when the independent variable value is specified. The regression equations for this study are as follows:

$$Y = 31.830 - 0.298X_1 - 0.178X_2 + 0.043X_3 + 2.444$$

which can be defined as below:

- Y = Level of text readability
- X<sub>1</sub> = Common and frequent words
- X<sub>2</sub> = Conjunction and punctuation
- X<sub>3</sub> = Average sentence length
- Constant = 31.830
- Error = 2.444

Once the regression model is developed, it is important to validate the suitability of the model. Model validity is the stability, reasonableness and ability of regression coefficients to generalize to other text samples in the population [21]. Regression model validation is an important step in the modeling process as it helps to evaluate the reliability of the model before it can be used in decision making [25].



Model validation can be seen on the adjusted R squared value. The adjusted R squared value gives an idea of how far the regression model can be generalized to the population. If the adjusted R squared values are equal to or close to the R squared values, this indicates that the model is compatible with the data set in the population [21]. As can be seen in table 4.16 below, the difference between the R-squared value and the adjusted R-squared value is 0.005 or 0.5% ( $0.446-0.441 = 0.005$ ). This shows that the model is compatible with the population. The reduction of the small adjusted R-squared value indicates that, if the model was derived from the population, it would account for the 0.5% variance reduction in the regression model produced.

**Table 7.** Summary of Model

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>
3	0.668 <sup>c</sup>	0.446	0.441

c.Independent Variables: Common and frequent words, Conjunction and punctuation, Average sentence length.

The suitability of a model to the data depends on the accuracy of the model in predicting and validating another or new dataset. This is the main purpose of model validation. Therefore, in addition to looking at the adjusted R-squared values, the researchers also used the Bootstrap approach in the SPSS software to ensure model validation accuracy. The Bootstrap method is an appropriate and more accurate method for testing the validity of the regression model without requiring large sample sizes [22, 23].

**Table 8.** Coefficients

Model	B	Standard error	Beta $\beta$	t	Sig.	95.0% Confidence Interval for B	
						Lower	Upper
Common and frequent words	-0.298	0.025	-0.522	-11.787	0.000	-0.348	-0.249
Conjunction and punctuation	-0.178	0.027	-0.283	-6.459	0.000	-0.232	-0.123
Average sentence length	0.043	0.019	0.097	2.246	0.025	0.005	0.080
Constant	31.830	2.444		13.026	0.000	27.022	36.638

Dependent Variable: Level of Text Readability

**Table 9.** Bootstrap for coefficients

Model	B	Bootstrap				
		Bias	Standard error	Sig. (2-tailed)	Bca 95% Confidence Interval	
					Lower	Upper
Common and frequent words	-0.298	0.000	0.020	0.000	-0.335	-0.260
Conjunction and punctuation	-0.178	-0.001	0.026	0.000	-0.226	-0.130
Average sentence length	0.043	0.001	0.017	0.014	0.009	0.079
Constant	31.830	-0.015	1.907	0.000	28.020	35.470

Bootstrap Results Based on 2000 Bootstrap samples

Based on tables 8 and 9, it can be seen that there is a similarity between the original stepwise regression model and the Bootstrap regression model based on 2000 re-samples. Similarity can also be seen through standard error values. For common and frequent words, the difference between the original model and the Bootstrap model is

0.005. The standard error values show values of 0.025 and 0.020, respectively. Further, the mean difference in standard error values for conjunction and punctuation in the original and Bootstrap models was 0.001, with the values of 0.027 and 0.026 respectively. Additionally, the mean difference in the standard error of the average sentence length in the original model and the Bootstrap model was 0.002 with values of 0.019 and 0.017 respectively.

The confidence interval values in the original model and the bias corrected acceleration confidence interval values for the Bootstrap models also show small differences. The lower and upper values of the common and frequent words in the original model were -0.348 and -0.249, while the lower and upper values in the Bootstrap model were -0.335 and -0.260. The differences between the lower and upper values in the original and the Bootstrap models for common and frequent words were 0.013 and 0.011, respectively. While the lower and upper values for the conjunction and punctuation in the original models were -0.232 and -0.123, respectively, while the lower and upper values in the Bootstrap models were -0.226 and -0.130. The differences between the lower and upper values in the original model and the Bootstrap model for conjunction and punctuation were 0.006 and 0.007, respectively. Furthermore, the lower and upper values of the average sentence length in the original model are 0.005 and 0.080, respectively, while the lower and upper values in the Bootstrap model are 0.009 and 0.079. The differences between the lower and upper values in the original model and the Bootstrap model for the average sentence length were 0.004 and 0.001, respectively.

The higher the similarity of the original regression model and the Bootstrap model, the higher the validity of the model [22, 26, 27]. This shows the stability between the regression model and the Bootstrap model for predicting the contribution of three variables which are; common and frequent words, conjunction and punctuation and average sentence length. Therefore, the regression model derived from the sample can be generalized to other data in the population.

Based on regression equations using the stepwise method, it can be identified that three independent variables were correlated and had an effect and contribution to the level of readability of the Arabic text. The three independent variables are common and frequent words, conjunction and punctuation as well as the average sentence length.

## 6 DISCUSSION

Based on the findings, the potential linguistic features that can influence the text readability consistently focus on complex sentences and simple sentences based on the frequency of their use in the text being studied. Theoretically, the word features also have the potential to affect the text readability in which the text being studied has a uniform frequency distribution in the use of conjunction and punctuation as well as the use of common and frequent words. Although the distribution of average sentence length is less uniform, it is however can influence the text readability in theory.

This finding is in line with [16, 3, 28] who studied the readability of 15 texts from Malaysian form four Arabic language (higher level) and *Balāghah* (Arabic rhetoric) textbooks, both of which were used in cloze test. They found that the use of complex sentences dominated the text with uniform use. Although common and frequent words also found to be dominating the text, their use, however, was not uniform.

However, the studies [7, 8] showed that the use of common and frequent words as well as the average sentence length contributed to the text readability. In contrast to [29] which showed that consistent use of complex sentences and abstract words in the text contributed to the readability of the text. However, a large number of samples and text used by native speakers were involved [7], whereas only small samples were used in [16, 28].

Based on previous readability studies, the previous researchers have found that linguistic features such as word elements that help readers' understanding and words that are intended for learning. New words were provided with keywords that

help define those words. In addition, the structure and type of sentence are often used as a measure of text readability such as sentence structure with explanations, examples and illustrations with the help of words that connects them to the content.

It can be concluded from all these findings that features that have a uniform use distribution have the potential to contribute to predicting the text readability. Likewise, in providing textbooks to non-native learners, the uniformity of linguistic features that could potentially influence the text readability should be tailored to the readers' level of competency according to the level of learning, not just based on the teaching content that should be followed by the level of learning. The linguistic features used in delivering content should be in accordance with the level of competency of the target readers.

This is in line with the recommendations in Klare Reader Performance Model [30] where it is important for the authors to take into account the reader's competence in providing them with appropriate text to ensure effective interaction between non-native language learners and Arabic textbooks. The text is used not only to learn the content of the subject, but also to learn Arabic as a foreign or second language.

The findings of this study also show the significant relationships between linguistic features of common and frequent words, conjunction and punctuation, complex sentences and average sentence length with the level of text readability. This indicates that the higher the level, the lower the frequency of using common and frequent words. Similarly, the higher the level, the lower the use of conjunction and punctuation. In terms of sentence, the higher the level, the lower the use of complex sentences in the text. For the average sentence length, the higher the level, the higher the average sentence length. The features that have significant relationships to the level of text readability can potentially contribute to the level of text readability.

In this study, only four out of seven linguistic feature variables had significant correlations with level of text readability variable, which are common and frequent words, conjunction and punctuation, average sentence length and complex sentence. Accordingly, these four variables were included in the multiple regression analysis.

Overall, the findings show only three linguistic features that highly contribute to the level of Arabic texts readability, which are, common and frequent words, conjunction and punctuation as well as average sentence length. The three variables contributed 44.6 percent to the variants in the level of text readability. The results showed that two predictor variables had negative effects, which are common and frequent words and conjunction and punctuation, while the variable of average sentence length had a positive effect. Based on this finding, it can be concluded that these three variables are able to contribute or to predict the level of Arabic text readability used by non-native students.

The linguistic features found to contribute to the level of text readability need to be adapted to the target readers' competency level which is in accordance with the Klare Reader Performance Model [30]. The readers' competency level covers a variety of aspects, which are, the reader's prior knowledge of the content presented in the reading material, reading skills, readers' language skills and the reader's thinking maturity. It depends on how the author controls the content density, adjusting the linguistic features such as using difficult or technical words as well as common and frequent words. Similarly, sentence structures need to be adjusted, for instance, long sentences, simple sentences and complex sentence.

In addition, the level of text readability needs to vary according to the readers' competency and the content density conveyed to the reader. The linguistic features of the text should be tailored to the content or information conveyed to the reader, whether the content is new information or continuation of what is already known, whether the content conveyed is interesting or fulfilling the readers' reading needs, whether the nature of the content corresponds to the readers' intellectual level and whether the maturity of the content consistent with the reader's thinking maturity.

However, there is a possibility for other factors that were not considered in this study to contribute to the readability of Arabic texts especially for non-native

learners. The results of this study are in line with the empirical evidence from previous studies that linguistic features such as common and frequent words, conjunction and punctuation as well as sentence lengths contribute and have their own influence on the level of Arabic text readability.

In the effort to provide text for non-native learners, attention should be given to aspects of linguistic features that contribute to the text readability such as the use of common and frequent words, the use of conjunction and punctuation to show the connection between ideas in sentences and between sentences as well as explanations of a concept or word. The average sentence length should also be noted even though the contribution of these features is mediocre.

The findings of this study generally support the appropriateness of reading theory and provide input to the cognitive behavior of reading for Arabic language learners as a foreign or second language. The study also strengthens the theory that reading is an interactive process that refers to the interaction between readers' knowledge related to the knowledge of letters, words, syntax, lexical, semantics and information of linguistic features of the text. Therefore, the linguistic features in the text should be given attention to match the development of students' language proficiency, especially non-native speakers. In addition, the section involved in the preparation of textbooks and reading materials in Arabic particularly should take into account the frequency of use of linguistic features appropriate to the level of Arabic language proficiency of students in Malaysia.

## **7 CONCLUSION**

Based on the findings, it can be concluded that texts for non-native learners should be carefully organized according to the level of learning based on the linguistic features that are found to contribute to the level of text readability. This is an effort to take into account the appropriate level of students' language mastery to ensure the effectiveness of reading process as well as encourage students to improve their Arabic proficiency. Further, the advances in information technology and use of internet in these recent years imply the need for inventing an updated approach to assessing Arabic reading materials readability. Taking this into consideration, the formula presented in this study is a modern upgrade from the conventional readability formulas. This study thus, contributes to the body of knowledge by providing a benchmark in measuring the readability of Arabic reading materials for non-native students. Using mathematical equation, this formula can be interpreted, allowing it to be easily integrated into any readability tools.

It can also be confirmed from this study that linguistic features of common and frequent words, conjunction and punctuation as well as the average sentence length contribute to the level of Arabic textbooks readability in the religious curriculum in Malaysia. Additionally, it is found that linguistic features that theoretically contribute to the readability of Arabic texts are not uniform across student levels and do not take into account the students' language mastery. Such linguistic features such as the frequency of abstract words, the use of simple sentences, complex sentences and noun phrases have their own strengths to contribute to the level of Arabic texts readability. However, some features of words and sentences help in understanding such as the use of common and frequent words, conjunction and punctuation that serve as indicators of the relationship between ideas and other ideas and the average sentence length. These linguistic features contribute to the level of text readability and it is found that the level of frequency and density in the text is uniform.

It is suggested for future studies to employ the findings of this study as a benchmark to research the development of readability methods or formulas for estimating text readability among non-native students.

## References

1. Kamarulzaman Abdul Ghani. 2011. Kebolehbacaan Buku Teks Bahasa Arab Tinggi Berasaskan Ujian Kloz dalam Kalangan Pelajar di SMKA [Readability of Arabic Language (Higher Level) Textbooks Based on Cloze Test among SMKA students]. *GEMA Online Journal of Language Studies* Volume 11 (2): 53-66.
2. Pintrich, P. R. 1999. The Role of Motivation in Promoting and Sustaining Self-Regulated Learning. *International Journal of Educational Research* 31: 459-470.
3. Kamarulzaman Abdul Ghani, Ahmad Sabri Noh & Nik Mohd Rahimi Nik Yusoff. 2014. Linguistic Features for Development of Arabic Text Readability Formula in Malaysia: A Preliminary Study. *Middle-East Journal of Scientific Research* 19(3): 319-331.
4. Al-Khalifa, H. S. & Al-Ajlan, A. A. 2010. Automatic Readability Measurements of The Arabic Text: An Exploratory Study. *The Arabian Journal for Science and Engineering* 35(2C): 103 - 124.
5. Aula A. 2004. Enhancing the Readability of Search Result Summaries, in *Proceedings of HCI Conference, Design for Life*, Leeds, UK, vol. 2, pp. 1-4.
6. Basu S. 2010. *8 Readability Web Tools to Test Your Writing Quality*. [online] Available from: <http://www.makeuseof.com/tag/writing-reader-friendly-check-8-readability-testing-webtools/> [Accessed 21 June 2020].
7. Dawood, B.A.K. 1977. The Relationship Between Readability and Selected Language Variables. Tesis Unpublished Master Thesis (In Arabic), Iraq Baghdad University.
8. Al-Heeti, K.N. 1984. Judgment Analysis Technique Applied To Readability Prediction of Arabic Reading Material. Tesis PhD, University of Northern Colorado.
9. Mat Daud, N., Hassan, H. & Abdul Aziz, N. 2013. A Corpus-Based Readability Formula for Estimate of Arabic Texts Reading Difficulty. *World Applied Sciences Journal* 21:168-173.
10. Al-Tamimi, A.-K., Jaradat, M., Aljarrah, N. & Ghanem, S. 2014. AARI: Automatic Arabic Readability Index. *The International Arab Journal of Information Technology*, 11(4): 370-3.
11. Kamarulzaman Abdul Ghani, Hassan Basri Awang Mat Dahan & Nik Mohd Rahimi Nik Yusoff. 2009b. Kebolehbacaan Buku Teks dan Hubungannya Dengan Persembahan Teknikal: Kajian terhadap buku teks Bahasa Arab Tinggi Tingkatan Empat [Textbook Readability and Its Relation to Technical Presentations: A Study on Arabic Language (Higher level) Form Four Textbook]. *Jurnal ISLAMIYYAT* Jilid 31, 3-27.
12. Al-Ajlan, A.A., Al-Khalifa, H.S. & Al-Salman, A.M.S. 2008. Towards The Development of An Automatic Readability Measurements for Arabic Language. *Third International Conference on Digital Information Management*, hlm. 506-511.78.
13. Kamarulzaman Abdul Ghani & Hassan Basri Awang Mat Dahan. 2009. Tahap Kebolehbacaan Buku Teks Bahasa Arab Tinggi Tingkatan Empat dan Hubungannya dengan Motivasi Murid SMKA [Arabic Language (Higher level) Form Four Textbook readability level and its Relation to SMKA Students' Motivation]. *Jurnal Pendidikan* Jilid 34: 55-72.
14. Kamarulzaman Abdul Ghani, Hassan Basri Awang Mat Dahan & Nik Mohd Rahimi Nik Yusoff. 2009a. Penggunaan Ujian Kloz dalam Mengukur Kebolehbacaan Teks Arab untuk Pembaca Bukan Arab di Malaysia: Satu Kajian Rintis [Using Cloze Test in Measuring Arabic Text Readability for Non-Arabic Readers in Malaysia: A Pilot Study]. *Journal of Islamic and Arabic Education* Issue 1, Vol 2: 15-30.
15. Kamarulzaman Abdul Ghani, Sabri Che Noh & Nik Mohd Rahimi Nik Yusoff. 2017. Ciri-Ciri Linguistik dalam Buku Teks Berbahasa Arab dan Hubungannya dengan Tahap Kebolehbacaan Teks di Malaysia [Linguistic Features in Arabic Textbooks and their Relationship to the Level of Text Readability in Malaysia]. *GEMA Online® Journal of Language Studies* Volume 17(3), August 2017.
16. Kamarulzaman Abdul Ghani. 2010. Kebolehbacaan Buku Teks Bahasa Arab Tinggi Tingkatan Empat Sekolah Menengah Kebangsaan Agama [Readability of Arabic Language (Higher Level) Form Four Religious Secondary Schools Textbooks]. Tesis PhD, Universiti Malaya.
17. Gall, M. D., Gall, J. P. & Borg, W. R. 2003. *Educational Research: An Introduction*. 7th Ed. Boston: Allyn and Bacon.
18. Abdul Rahman Chik. 1994. *Ma'āyir al-Tahakkum fī al-Alfaz fī Ta'lim al-ʿArabiyyah lil Mujtama'āt al-Islāmiyyah fī Janub Syarqi Asia*. Tesis PhD, Universiti Iskandariyyah.
19. Tabachnick, B.G. & Fidell, L.S. 2007. *Using Multivariate Statistics*. Fifth Edition. USA: Pearson Education, Inc.

20. Pallant, J. 2011. *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS*. 4th Ed. Crows Nest, NSW: Allen & Unwin.
21. Field, A. 2005. *Discovering Statistics Using SPSS*. London: Sage Publications.
22. Efron, B. & Tibshirani, R.J. 1993. *An Introduction to The Bootstrap*. New York: Chapman & Hall.
23. Wood, M. 2004. Statistical Inference Using Bootstrap Confidence Intervals. *Significance* 1(4): 180-182.
24. Cohen, J.W. (1988). *Statistical Power Analysis for The Behavioral Sciences* (2<sup>nd</sup> edition). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
25. Stevens, J.P. 2009. *Applied Multivariate Statistics for The Social Sciences*. Fifth Edition. New York: Routledge Taylor & Francis Group.
26. Oredein, A., Olatayo, T. & Loyinmi, A. 2011. On Validating Regression Models with *Bootstraps* and Data Splitting Techniques. *Global Journal of Science Frontier Research* 11(6).
27. Scalon, J. D., Freire, S. M. & Cunha, T. A. 1998. Validation of Models for Predicting The Use of Health Technologies. *Medical decision making* 18(3): 311-319.
28. Zulazhan Ab. Halim. 2012. Kebolehbacaan Buku Teks Balāghah Sijil Tinggi Agama Malaysia di Sekolah-Sekolah Menengah Agama Negeri [Readability of the Malaysian Religious High Certificate *Balāghah* Textbook in State Religious Secondary Schools]. Tesis PhD, Universiti Kebangsaan Malaysia.
29. Al-Durayhim, N.A. 1998. Mada tawaffur khasa'is al-inqira'iyyah wa al-nas al-jayyid al-'i'dad fi nusus kitab al-qira'ah fi al-lughah al-injiliziyyah li al-saf al-thani li al-banat wa mada ma'rifah al-mu'allimat bi hazih al-khasa'is. Tesis Sarjana, King Saud University, Riyadh.
30. Klare, G.R. 1988. The formative years. Dlm. Zakaluk, B.L. & Samuels, S.J. (pnyt.). *Readability: Its Past, Present and Future*. hlm. 14-34. Newark, DE: International Reading Association.