International Journal of Evaluation and Research in Education (IJERE)
Vol.4, No.2, June 2015, pp. 45~53
ISSN: 2252-8822

Validation Study of Waray Text Readability Instrument

Voltaire Q. Oyzon[1], Juven B. Corrales[2], Wilfredo M. Estardo, Jr.[2]
[1] College of Arts & Sciences, Leyte Normal University, Philippines
[2] Graduate School, Leyte Normal University, Philippines

Article Info

Article history:

Received Feb 16, 2015
Revised May 14, 2015
Accepted May 21 , 2015

Keyword:

Assessment
MTBMLE
Text Readability
Validation study
Waray Text Readability Instrument

ABSTRACT

In 2012 the Leyte Normal University developed a computer software—modelled after the Spache Readability Formula (1953) made for English—made to help rank texts that can is used by teachers or research groups on selecting appropriate reading materials to support the DepEd's MTB-MLE program in Region VIII, in the Philippines. However, "several experiments have already established that existing readability measures in English cannot directly be used to compute readability of other languages." To validate the Waray Text Readability Instrument (WTRI) formula, 15 stories were rated by 24 randomly selected teachers from two elementary schools in Tacloban City. The WTRI software uses two factors in determining readability, namely: (a) sentence length and (b) frequency of commonly occurring words. The teachers' task is to read the given text and rate the grade level of each text by considering these three factors: (1) frequency of commonly used words; (2) sentence length; and, (3) total number of words. The data gathered was compared with the WTRI's ratings of the same texts. Statistical testing was done to determine if there is a significant difference between the teachers' rating of the texts and the WTRI's ratings. As a result, there was no significant difference between the software's grade level ratings and that of the teachers'. It implied that the WTRI's calculation is valid.

Corresponding Author:

Voltaire Q. Oyzon,
College of Arts & Sciences,
Leyte Normal University,
P. Paterno St, Tacloban City 6500, Philippines
Email: v.oyzon@gmail.com

1. INTRODUCTION

A readability algorithm in readability software is used to help rank texts. Put simply, readability determines how easy it is to comprehend a text. Reference [1] established that longer sentences are harder to read than short sentences and longer words are harder to comprehend than short ones. Most readability formulas calculate the average length of sentences and the average number of syllables to give a readability score.

Readability algorithms are highly accurate measures of comprehension level, but the original formulas were developed for English, an Anglo-Saxon language with many single-syllable words. In contrast, Waray has very few single syllable words. Almost no nouns, verbs, or adjectives are monosyllablic. Further, Waray-Waray has very few prepositions, in comparison to the many monosyllabic prepositions in English. In short, Waray is polysyllabic, and furthermore, its grammar is agglutinative. Prefixes and suffixes change parts of speech. Other parts of speech are formed through affixes:

- sumat (tell) --> mag**sumat** (telling), magsu**sumat** (will tell), su**sumat**on (tales)
- ada (there) --> may**ada** (has/have), magkamay**ada** (let there be)

Therefore, standard readability index like the Flesch-Kincaid formula would categorize Waray language texts as much more difficult to comprehend, simply because they have more syllables. Furthermore,

syllabification is different in Waray than in English. Vowels are never combined into one syllable (in English, "too" is one syllable; in Waray, "tuod" is two syllables).

The Waray Readability Formula

Leyte Normal University made a readability formula tailored for the Waray language: (a) sentence length and (b) frequency of common words determine readabilty; syllable length is disregarded. Knowing the frequency of words as they are used in written and oral communication provided the best means of inferring the likelihood that a word would be encountered by a reader and thus become part of that individual's receptive vocabulary [2]. These criteria are based on researches by [1],[3]-[5].

In the Philippines, no studies and effort have been made in the development of readability formula for the non-Tagalog languages until the passage of MTBMLE into a law 2012. In 2012, Leyte Normal University, with the help of John Mark Fullmer of Austin College, initiated the development of readability formula or, software for Waray language [6]. However, this readability formula is patterned after Spache Readability Formula made for English. Therefore, adjustments have to be made to suit the idiosyncrasies of the Waray language. This study aims to validate this Waray Text Readability Instrument.

This study aims to investigate the following questions: (a) What is the difference between the WTRI (or, software's) grade level assessment and the teachers' grade level assessment of the fifteen selected Waray texts? Is there a significant difference between the two? ; and (b) What is the appropriate formula for the software? Can we formulate a new algorithm from the gathered data? This study cross-validate the Waray Text Readability Instrument and seeks for an efficient text readability formula tailored for Waray language to support the MTB-MLE program of the DepEd. To answer these questions, we conducted an experiment. Selected elementary teachers were asked to read fifteen (15) selected Waray texts, and rate what grade level each text suits by considering these three (3) factors: (a) word content; (b) sentence length; and (c) frequency of difficult words. With those grade level ratings, we compared the data with the WTRI's grade level ratings and we administered statistical testing to determine if there is a significant difference between the two.

2. RESEARCH METHOD

The participants were the teachers from two different schools under Tacloban City Division, Tacloban City, Leyte. We selected San Fernando Central School and Panalaron Elementary School. The school principals randomly picked the twelve (12) teachers in their respective schools having a total of twenty-four (24) teachers as our respondents. The twelve (12) teachers from San Fernando Central School (SCS) were the first set of respondents and the twelve (12) teachers from Panalaron Elementary School (PES) were the second set of our respondents.

The material used in the study was taken from the software collections of Waray texts submitted by different authors. We randomly picked 15 samples from the corpus with different grade levels and genres. Out of those fifteen (15) samples, five (5) of them were graded two (2), four (4) were graded three (3), three (3) were graded four (4), and three (3) were graded five (5). Also, out of those fifteen (15) again, five (5) of them were essays, five (5) were poems, four (4) were stories, and one (1) was uncategorized.

This study used the experimental research design by comparing the means of the software and the gathered data. We used two statistical test; (a) t-test for two dependent samples; and One-Way Analysis of Variance to answer our question if there is a significant difference between those results. We used t-test for two dependent samples to compare the mean of the software and the mean of the teachers. Also we used the One-Way Analysis of Variance to compare the means of the three groups. This will tell if there is a significant difference between the software and teacher of SCS, software and teacher of PES, and teacher of SCS from teacher of PES.

3. RESULTS AND ANALYSIS

Table 1 below shows the different ratings (grade levels) of selected Waray texts provided by the elementary teachers of San Fernando Central School (SFCS) and Panalaron Central School (PCS).

Table 1. WTRI vs Teachers' Grade Level Ratings of Texts

No.	Title of Waray Story	Genre	WTRI's Rating	SFCS	PCS	AT Rating
1	An Karabaw	Story	3.1	1	3	2
2	Sigbin	Essay	4.4	4	4	4
3	A Buong Nga Lahog	Story	5.2	3	5.5	4.25
4	An Akon La Nahihinumduman	Essay	4.7	6	6	6
5	Amo La Gihapon	Essay	4.6	5	5	5
6	Harupihap	Story	6.6	6	6	6
7	An Aswang	Essay	5.4	4	4	4
8	An Madulom Nga Kagab-ihon	Essay	6.3	6	6	6
9a	Ako Anak Hin OFW	Poem	3.1	3	3	3
9b	Situwasyon	Poem	2.8	3	2	2.5
10a	O Bulan	Poem	3.7	3	2	2.5
10b	Naghihinglaw	Uncategorize	2.0	4	2	3
11a	An Ngaran Nga Nanay	Poem	3.5	3.5	3.5	3.5
11b	Dagaw	Poem	4.0	3.5	4.5	4
12	A Bayod	Story	4.8	4.5	4	4.25

For convenience, we utilized the following abbreviations:

WTRI = Waray Text Readability Instrument
SFCS = Results from San Fernando Central School teachers
PCS = Results from Panalaron Central School teachers
AT Rating=Average Teachers' rating; **(**SFCS + PCS) / 2 or, Average of two schools

In section 3.1, the WTRI's rating and the average teachers' rating of the texts were subjected to a t-test to find out if there is a significant difference. In section B t-test were applied among the three: (a) software's ratings of the texts; (b) ratings of the texts by San Fernando Central School teachers; and (c) ratings of the texts by Panalaron Elementary School teachers.

3.1. Hypothesis Testing

A. Test if there is a significant difference between the software and teachers' gathered data. Use alpha = 0.05

Table 2. WTRI vs Mean Teachers' Grade Level Ratings of Texts

WTRI ratings	3.1	4.4	5.2	4.7	4.6	6.6	5.4	6.3	3.1	2.8	3.7	2.0	3.5	4.0	4.8
AT Ratings	2.0	4.0	4.25	6.0	5.0	6.0	4.0	6.0	3.0	2.5	2.5	3.0	3.5	4.0	4.25

Table 2 illustrates the different grade level ratings of the texts (see on table 1) between the WTRI and the average teachers' ratings (SFCS +PCS).

1. H_0: There is no significant difference between the software and the teachers' result.
2. H_a: There is a significant difference between the software and the teachers' result.
3. $\alpha = 0.05$
4. **Statistical Test**: t-test
5. **Critical Value**: $t_{\frac{0.05}{2},14} = \pm 2.145$; **Critical Region**: $t_c < -2.145$ or $t_c > 2.145$

Table 3. Mean, Standard Deviation of WTRI and AT Grade Level Ratings

	WTRI	AT	d	d^2
1	3.1	2.0	1.10	1.21
2	4.4	4.0	0.40	0.16
3	5.2	4.25	0.95	0.90
4	4.7	6.0	-1.30	1.69
5	4.6	5.0	-0.40	0.16
6	6.6	6.0	0.60	0.36
7	5.4	4.0	1.40	1.96
8	6.3	6.0	0.30	0.09
9	3.1	3.0	0.10	0.01
10	2.8	2.5	0.30	0.09
11	3.7	2.5	1.20	1.44
12	2.0	3.0	-1.0	1.00
13	3.5	3.5	0.00	0.00
14	4.0	4.0	0.00	0.00
15	4.8	4.25	0.55	0.30
			$\sum d$ = 4.20	**$\sum d^2$ = 9.38**

Table 3 shows, the first column (numbers 1 to 15) the texts mentioned in table 1. The second column indicates the Waray Text Readability Instruments' (WTRI) grade level ratings. The third column signifies the average mean of each text as rated by the teachers (SFCS +PCS). Fourth column points the difference between the WTRI rating from AT's mean rating (WTRI minus AT). To calculate, for instance, the difference of mean for text 1, get the difference between the WTRI ratings and average teachers' rating (AT), which is 1.10 (see fourth colum **d)**; text 2 has a difference of 0.40, and so on. Add all these diffrences, and we got a total of 4.20. We will use this value later in the computation of variance. Lastly, the fifth column (d^2) indicates the squared of the values in the fourth column **d.** For example, for text 1, d= 1.10, the value 1.10 is squared which is 1.21. The same mathematical process is applied to texts 2 to 15. The total of all the d^2, will be needed in our computation of standard deviation below.

6. **Computation**:

$$\bar{d} = \frac{\sum d}{n} = \frac{4.20}{15} = \mathbf{0.28}\ ; \qquad s_d = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}} = \sqrt{\frac{9.38 - \frac{(4.20)^2}{15}}{15-1}} = \mathbf{0.77}$$

$$t_c = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{0.28}{0.77/\sqrt{15}} = \mathbf{1.408}$$

1. **Decision**: Do not reject H_o, since 1.408 < 2.145.
2. **Conclusion**: There is sufficient evidence that there is no significant difference between the software and teachers' grade level ratings of the texts.

Table 4 shows the different reading level ratings of the 12 texts given by the Waray Text Readability Instrument (WTRI), by the teachers of San Fernando Central School, and by the teachers of Panalaron Central School. Below is our computation of the means and variances.

Table 4. WTRI vs. SFCS vs. PCS Grade Level Ratings of Texts

WRTI rating	Teachers from SFCS rating	Teachers from PCS ratings
3.1	1	3
4.4	4	4
5.2	3	5.5
4.7	6	6
4.6	5	5
6.6	6	6
5.4	4	4
6.3	6	6
3.1	3	3
2.8	3	2
3.7	3	2
2.0	4	2
3.5	3.5	3.5
4.0	3.5	4.5
4.8	4.5	4

B. Test if there is significant difference among the three: (a) software; (b) San Fernando Central School; and (c) Panalaron Elementary School.

1. H_o: All means are equal
2. H_a: Not all means are equal
3. $\alpha = 0.05$
4. **Statistical Test**: F-test (One-way ANOVA)
5. **Critical Value**: $F_{0.05,\ 2,42} = 3.23$; **Critical Region**: $F_c > 3.23$

Table 5 indicates the WTRI's mean which at 4.28; variance at 1.67; SFCS's mean is at 3.97; variance at 1.91; and PCS's mean is at 4.03; variance is at 2.12.

6. **Computations**:
   a. Compute the means and variances of the three groups.

Table 5. WTRI vs. SFCS vs. PCS Mean and Variance

WTRI	SFCS	PCS
Mean = 4.28 Variance = 1.67	Mean = 3.97 Variance = 1.91	Mean = 4.03 Variance = 2.12

   b. Find the grand mean:

$$\bar{X}_{GM} = \frac{\sum X}{N} = \frac{3.1 + 4.4 + 5.2 + \cdots + 3.5 + 4.5 + 4}{15} = 4.09$$

   c. Find the between – group variance

$$S_B^2 = \frac{\sum n_i (\bar{X}_i - \bar{X}_{GM})^2}{k-1}$$
$$= \frac{15(4.28-4.09)^2 + 15(3.97-4.09)^2 + 15(4.03-4.09)^2}{3-1}$$
$$= \frac{0.81}{2} = \mathbf{0.406}$$

   d. Find the within – group variance

$$s_W^2 = \frac{\sum(n_i - 1)\, s_i^2}{\sum(n_i - 1)}$$
$$= \frac{(15-1)(1.67)+(15-1)(1.91)+(15-1)(2.12)}{42}$$
$$= \frac{79.8}{42} = \mathbf{1.9}$$

   e. Find F_c

$$F_c = \frac{S_B^2}{S_W^2} = \frac{0.406}{1.9} = 0.214$$

Table 6. Variation between WTRI & AT/ Within AT

Source of Variation	Sum of Squares	df	Mean Square	F – Value	F Critical
Between WTRI & AT	0.812	2	0.406	0.214	3.23
Within AT	79.8	42	1.9		
Total	80.61	44			

7. **Decision:** Do not reject H_o, since $F_c < 3.23$

Table 6 shows the variation between WTRI ratings of the texts and the average of all the teachers' ratings (AT). Since the F-value is less than 3.23, this implies that the software, SCS teachers and PES teachers have no significant difference on ranking the texts. . In other words, there is sufficient evidence that all means are equal.

3.2. The Number of Words in a Text

"The more the words in a text, the higher the grade level it ranks," most of the respondents say. In the process of our data gathering, majority of the teachers suggested that the number of words in the text should be factored in in the validation of WTRI. We accepted the challenge of the teachers, and we formulated a new algorithm, an alternative to the existing WTRI. And we tested this new algorithm results with the WTRI ratings of the same texts (see table 7 below).

Let **A** be the total number of words
Let **B** be the percentage of not frequently occurring words

Table 7. Toral No. of Words vis-à-vis Percenatage of Non-Frequent Words

A \ B	10	20	30	40	50	60	
100	1	1.5	2	2.5	3	3.5	Grade Levels
200	1.5	2	2.5	2.0	3.5	4	
300	2	2.5	3	3.5	4	4.5	
400	2.5	3	3.5	4	4.5	5	
500	3	3.5	4	4.5	5	5.5	
600	3.5	4	4.5	5	5.5	6	

Table 7 simple suggests that for a story with 100 words and has 10% of its vocabulary is non-frequently occurring words; it should have a grade level rating of 1. Meaning, this text is suitable for grade 1. On the other hand, if a text has 600 total words, and has 60% of its vocabulary is non-frequent; it must have a rating level of grade 6. Meaning, this story is suited for grade 6.

In order to do this we need to formulate new new algorithm, an alternative to the existing WTRI.

Solve the numerical coefficient of the two factors

- Let x be the numerical coefficient of the total number of words
- Let y be the numerical coefficient of the percentage of not frequently occurring words

Based on the table, we can have,

$$100x + 10y = 1$$
$$200x + 10y = 1.5$$

Using this two equations, solve for x and y.
Sol.

$$\begin{array}{ll} 100x + 10y - 1 = 0 & \textbf{(1)} \\ -\ (200x + 10y - 1.5 = 0) & \textbf{(2)} \\ \hline -100x + 0 + 1 = 0 & \textbf{(3)} \end{array}$$

Solve for x using equation **(3).**

$$-100x + 1 = 0$$
$$1 = 100x$$
$$x = 1/100$$
$$\mathbf{x = 0.005}$$

Now solve for y using the value of x and equation **(1)** or **(2).**
Using the equation **(1),**

$$100(0.005) + 10y - 1 = 0$$
$$0.5 + 10y - 1 = 0$$
$$10y = 0.5$$

y = 0.5/10 or **0.05**

Aside from two factors above, sentence length also one of the factors that need to consider. Shorter sentences indicate simpler reading level. Very simple texts have less than 10 words per sentence. Difficult texts can have more than 20 [2].

Based on the given interval, **0.1** is the best fit numerical coefficient for the sentence length.

Proof:

Let z be the numerical coefficient of the average sentence length.
Suppose the average sentence length is 10, then it will be considered easy (grade I)
Now we can have this equation,

$10z = 1$

Then, $z = 1/10$ or $z =$ **0.1**

Therefore, 0.1 is the numerical coefficient for the sentence length.

Plus the constant as a balancer of the equation which is **0.87**.

Now the final formula is,

Grade Level = (0.005 x Total Number of Words) + (0.05 x Percentage of not frequently occurring word) + (0.1 x Average Sentence Length) + 0.87

Or

GL = **A**x + **B**y + **C**z + **0.87**

where **A** be the Total Number of Words
B be the Percentage of not frequently occurring word
C be the Average Sentence Length
x = 0.005
y = 0.05
z = 0.1

3.3. Comparison of the Results

This time we compared the grade level ratings of the texts by WTRI vs. SFCS vs. PCS vs. the new algorithm (New F). In table 8 below, the numbers in red indicate the extreme values or those values far away from the other values.

Table 8. WRTI vs. SFCS vs. PCS vs. New Formula Grade Level Ratings of Texts

No.	Title	WTRI	SFCS	PCS	New F.
1	An Karabaw	3.1	1	3	3.57
2	Sigbin	4.4	4	4	4.16
3	A Buong Nga Lahog	5.2	3	5.5	5.66
4	An Akon La Nahihinumduman	4.7	6	6	6.39
5	Amo La Gihapon	4.6	5	5	5.51
6	Harupihap	6.6	6	6	8.05
7	An Aswang	5.4	4	4	5.47
8	An Madulom Nga Kagab-ihon	6.3	6	6	8.01
9	Ako Anak Hin OFW	3.1	3	3	2.94
10	Situwasyon	2.8	3	2	2.45
11	O Bulan	3.7	3	2	2.96
12	Naghihinglaw	2.0	4	2	1.99
13	An Ngaran Nga Nanay	3.5	3.5	3.5	2.87
14	Dagaw	4.0	3.5	4.5	3.08
15	A Bayod	4.8	4.5	4	4.16

Further test was administered and resulted to the non-rejection of the null hypothesis, which means, there is no significant difference between the existing WTRI formula and that of the new formula, where the total number of words in a text was factored in, as suggested by the respondent/teachers.

One explanation to this would be from [7] who posits that "the comprehensibility or difficulty of a message is dominated by the familiarity of the semantic units and by the complexity of the syntactic structures used in constructing the message *not by the number of words*" (emphasis added).

4. CONCLUSION

The evidence in this study led to the following conclusions: (1) From the first phase of the study, determining the difference between the WTRI/software's and the teachers' grade level ratings, the conclusion reached is that the existing Waray Text Readability Formula has a valid result on calculating the readability of the particular text as determined by the teachers' analysis on ranking the text. (2) On the second phase, determining the difference among the three: (a) WTRI; (b) San Fernando Central School teachers' rating; and (c) Panalaron Central School teachers' rating, resulted in not rejecting the null hypothesis. This implies that, all in all, the San Fernando Central School teachers' ratings have no significant difference when compared to the ratings given by the teachers of Panalaron Central School, as well as with that of the software (WTRI). (3) Lastly, the formulation of the new algorithm was the product of the suggestions of our respondents that the number of words must be included as one of the factors. The total number of words in a text was disregarded since they were particular on the syntactic and semantic components to determine the readability of the text. However, the new formula has a valid calculation and there was no significant difference exists between the new formula and the WTRI current formula.

ACKNOWLEDGEMENTS

The following people willingly extended help in one way or another during the conduct of this project: John Mark Fullmer, Lorena M. Ripalda, Ms. Gretel Cadiong, MTBMLE Coordinator for DepEd Tacloban City Division, the principal and teachers of Panalaron Central Elementary School, and the principal and teachers of San Fernando Central Elementary School.

REFERENCES

[1] Flesch R., "A new readability yardstick", *Journal of Applied Psychology*, vol/issue: 32(3), pp. 221-233, 1948. [http://dx.doi.org/10.1037/h0057532].
[2] Carroll JB., Davies P., Richman B., "The American Heritage Word Frequency Book", New York: American Heritage, 1971.
[3] Dale E., Chall J., "A formula for predicting readability", *Educational Research Bulletin*, vol. 27, 1948.
[4] Spache G., "A new readability formula for primary-grade reading materials", *Elementary school journal*, vol. 53, pp. 410-413, 1953.
[5] MetaMetrics, Inc., "The Lexicile Framework for Reading: Technical Report", 2007. Retrieved from < https://www.azed.gov/wp.../LFforReadingTechnicalReport_042007.pdf>
[6] Oyzon V., Fullmer JM., **"**Mother-Tongue Based Multilingual Education (MTBMLE) Initiatives in Region 8", *International Journal of Evaluation and Research in Education (IJERE)*, vol/issue: 3(1), pp. 53-65, 2014.
[7] Carver RP., "Measuring the primary effect of reading: Reading storage technique, understanding judgments and cloze", *Journal of Reading Behavior*, vol. 6, pp. 249–274, 1974.

BIOGRAPHIES OF AUTHORS

Voltaire Q. Oyzon is a member of Asia-Pacific Writers and Translators Association. He teaches literature and social sciences courses at the Leyte Normal University. In collaboration with others, he has edited books such as, "Syahan nga Hiwat: An Namagdaog han Syahan nga Pasidungog Eduardo A. Makabenta, Sr. Para han Siday (2010, with Michael Carlo C. Villas & Jose N. Lianza), published by Sinirangan Culture & Arts Development Center: Leyte Normal University, Tacloban City; "An Bag-o nga Ortograpiya han Winaray" (2012, with Ricardo Ma. D. Nolasco and Firie Jill Ramos), "Syahan nga Usa Ka Yukot nga mga Pulong nga Agsob Gamiton ha Winaray: Pagpurulongan para han mga Magturutdo ha MTBMLE" (2013, with John Mark Fullmer and Evelyn C. Cruzada) and, "Learner's Classified Dictionary for Learning English-waray-Kana-Inabaknon-Tagalog Vocabularies" (2014, with John Mark Fullmer and Evelyn C. Cruzada) published by the Philippines Commission on Higher Education.

Juven B. Corrales is currently working on his Master's degree major in Mathematics at the Leyte Normal University Graduate School.

Wilfredo M. Estardo, Jr Corrales is currently working on his Master's degree major in Mathematics at the Leyte Normal University Graduate School.