
A Second Generation Readability Formula for Vietnamese

Author(s): Liem T. Nguyen and Alan B. Henkin

Source: *Journal of Reading*, Vol. 29, No. 3 (Dec., 1985), pp. 219-225

Published by: [Wiley](#) on behalf of the [International Reading Association](#)

Stable URL: <http://www.jstor.org/stable/40029662>

Accessed: 01-03-2016 09:17 UTC

REFERENCES

Linked references are available on JSTOR for this article:

http://www.jstor.org/stable/40029662?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Wiley and International Reading Association are collaborating with JSTOR to digitize, preserve and extend access to *Journal of Reading*.

<http://www.jstor.org>

Nguyen is currently a research associate at the University of Iowa, Iowa City, Iowa, working with the linguistics of Indochinese languages and the behavioral adjustment of refugees. He was previously Vice-Minister of Education for the Republic of Vietnam. Henkin teaches in the College of Education at the University of Iowa, designing and developing multilingual reading materials and related evaluation procedures.

A second generation readability formula for Vietnamese

Liem T. Nguyen
Alan B. Henkin

In 1982, we published a readability formula for Vietnamese that used average sentence length and average word length as predictors of the readability level of Vietnamese prose (Nguyen and Henkin, 1982). Those two variables accounted for 74% of the variance in readability, leaving another 26% unexplained. Because of this low coefficient of determination, we have reexamined the formula. This article presents our revised formula and discusses its validity and reliability.

Theoretical considerations

Studies on readability show that the two most highly loaded factors are vocabulary difficulty, which is probably the most important, and sentence length (Brinton and Danielson, 1958; Stolurow and Newman, 1959). The most commonly used indexes of vocabulary difficulty are word length, determined by computing the average number of syllables per hundred words, and word frequency or word familiarity, determined by finding the percentage of difficult or unfamiliar words in a passage based on an established list of easy words.

Our original readability formula, which we will refer to as Formula 1,

Table 1
Correlations among variables (N = 24)

| | Readability level | Word difficulty | Sentence length | Word length |
|-------------------|-------------------|-----------------|-----------------|-------------|
| Readability level | 1.000 | | | |
| Word difficulty | .923 | 1.000 | | |
| Sentence length | .722 | .603 | 1.000 | |
| Word length | .693 | .723 | .435 | 1.000 |

used word length as an index of vocabulary difficulty. However, we began to question the extent to which word length accounts for vocabulary difficulty in Vietnamese and found that word length accounted for about 73% of the variance in vocabulary difficulty (Nguyen and Henkin, 1981a). This low coefficient of determination may be explained by certain characteristics of the Vietnamese language.

Most Vietnamese words consist of one syllable; for example, *ăn* 'to eat,' *đi* 'to walk,' *nhà* 'house.' Some compound or hyphenated words have been formed by combining two or three single words; for example, *Hồng-thập-tù* 'red cross,' and *học-sinh* 'student.'

Sixty-five percent of Vietnamese compound words are of Chinese origin, commonly referred to as *tiếng Hán-Việt* 'Sino-Vietnamese words.' These constitute a large portion of Vietnamese scientific and technical vocabulary and are considered more difficult than common Vietnamese vocabulary.

The remaining 35% of compound words are purely Vietnamese in origin and do not seem to be more difficult than noncompound words. *Lạnh-lẽo* 'cold' is a Vietnamese compound word; it is longer than the noncompound word *lạnh* 'cold,' and the two are equally familiar and popular.

When we used word length as a predictor in establishing Formula 1, we did

not differentiate between Vietnamese and Sino-Vietnamese compound words. Hence, word length was a less than optimal index of word difficulty.

Since compound Sino-Vietnamese words are the most difficult part of Vietnamese vocabulary, we felt it would be preferable to use the percentage of Sino-Vietnamese words in a passage instead of word length as an index of vocabulary difficulty. The multiple regression equation takes the following form:

$$RL = B_0 + B_1WD + B_2SL$$

where RL (the criterion) is the judged readability level, WD (first predictor) is the percentage of compound Sino-Vietnamese words, and SL (second predictor) is the average sentence length of a passage.

Developing and evaluating the new formula

After selecting 24 passages of approximately 100 words each from Vietnamese textbooks, popular novels, and textbooks, we computed the judged readability level (RL), the percentage of difficult words in a passage (WD), the average sentence length (SL), and the average word length (WL) of the passages in the following way.

1. We gave the passages to 10 former Vietnamese teachers with specializations in Vietnamese language arts; 4 were elementary school teach-

Table 2
Multiple regression analysis

| Variable | R | R ² | B | Constant |
|-----------------|------|----------------|------|----------|
| Word difficulty | .923 | .852 | .266 | |
| Sentence length | .946 | .895 | .129 | 1.740 |

| Analysis of variance | | | | |
|----------------------|----|---------|---------|---------|
| Source | df | SS | MS | F |
| Regression | 2 | 236.497 | 118.249 | 89.768* |
| Residual | 21 | 27.663 | 1.317 | |

* $p < .01$

ers, 5 were secondary school teachers, and 1 was a college instructor. All were graduates of either the Saigon Normal School or the Faculty of Pedagogy (Saigon University) and had 5 or more years of teaching experience.

They judged the readability level of each passage on a rating scale that had 13 points corresponding to the different grade levels of the Vietnamese education system (Nguyen and Henkin, 1981b) as follows: 1-5 = elementary, 6-9 = secondary first cycle, 10-12 = secondary second cycle, 13 = college level.

We computed the judged readability level (RL) for each passage by averaging the rating scores of the 10 teachers. The RL mean score of the 24 passages was 7.40 and the standard deviation was 3.389.

2. Vocabulary difficulty (WD) was measured by computing the percentage of compound Sino-Vietnamese words in each passage using the following formula.

$$WD = \frac{\text{number of compound Sino-Vietnamese words}}{\text{total number of words}} \times 100$$

The mean vocabulary difficulty score for the 24 passages was 12.375 with a standard deviation of 9.771.

3. The average sentence length (SL)

for each passage was obtained by dividing the total number of words by the total number of sentences in the passage.

$$SL = \frac{\text{total number of words}}{\text{total number of sentences}}$$

The mean sentence length for the 24 passages was 18.399 with a standard deviation of 6.838.

4. The average word length (WL) in each passage is rendered by the following formula.

$$WL = \frac{\text{total number of letters}}{\text{total number of words}}$$

The average word length for the 24 passages was 4.615 with a standard deviation of .826.

5. Correlations among the variables RL, WD, SL, and WL are presented in Table 1. The highest correlation was between the judged readability level (RL) and vocabulary difficulty (WD) ($r = .923$). Thus, WD was chosen as the first independent variable to be entered into the multiple regression equation.

The second highest correlation was between readability level (RL) and sentence length (SL) ($r = .722$), and SL was chosen as the second predictor in the multiple regression equation. The

Table 3
Readability levels (grade levels, Vietnamese educational system)

| | Sentence length (average number of words) | | | | | | | | | | | | | | | | | | |
|---|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|--|--|--|
| | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | | | | |
| Word difficulty (percentage of total words) | 0 | 2.4 | 2.5 | 2.6 | 2.8 | 2.9 | 3.0 | 3.2 | 3.3 | 3.4 | 3.6 | 3.7 | 3.8 | 3.9 | 4.1 | 4.2 | | | |
| | 1 | 2.7 | 2.8 | 2.9 | 3.0 | 3.2 | 3.3 | 3.4 | 3.6 | 3.7 | 3.8 | 4.0 | 4.1 | 4.2 | 4.3 | 4.5 | | | |
| | 2 | 2.9 | 3.1 | 3.2 | 3.3 | 3.4 | 3.6 | 3.7 | 3.8 | 4.0 | 4.1 | 4.2 | 4.4 | 4.5 | 4.6 | 4.7 | | | |
| | 3 | 3.2 | 3.3 | 3.5 | 3.6 | 3.7 | 3.8 | 4.0 | 4.1 | 4.2 | 4.4 | 4.5 | 4.6 | 4.8 | 4.9 | 5.0 | | | |
| | 4 | 3.5 | 3.6 | 3.7 | 3.9 | 4.0 | 4.1 | 4.2 | 4.4 | 4.5 | 4.6 | 4.8 | 4.9 | 5.0 | 5.2 | 5.3 | | | |
| | 5 | 3.7 | 3.9 | 4.0 | 4.1 | 4.3 | 4.4 | 4.5 | 4.6 | 4.8 | 4.9 | 5.0 | 5.2 | 5.3 | 5.4 | 5.6 | | | |
| | 6 | 4.0 | 4.1 | 4.3 | 4.4 | 4.5 | 4.7 | 4.8 | 4.9 | 5.0 | 5.2 | 5.3 | 5.4 | 5.6 | 5.7 | 5.8 | | | |
| | 7 | 4.3 | 4.4 | 4.5 | 4.7 | 4.8 | 4.9 | 5.1 | 5.2 | 5.3 | 5.4 | 5.6 | 5.7 | 5.8 | 6.0 | 6.1 | | | |
| | 8 | 4.5 | 4.7 | 4.8 | 4.9 | 5.1 | 5.2 | 5.3 | 5.5 | 5.6 | 5.7 | 5.8 | 6.0 | 6.1 | 6.2 | 6.4 | | | |
| | 9 | 4.8 | 4.9 | 5.1 | 5.2 | 5.3 | 5.5 | 5.6 | 5.7 | 5.9 | 6.0 | 6.1 | 6.2 | 6.4 | 6.5 | 6.6 | | | |
| | 10 | 5.1 | 5.2 | 5.3 | 5.5 | 5.6 | 5.7 | 5.9 | 6.0 | 6.1 | 6.3 | 6.4 | 6.5 | 6.6 | 6.8 | 6.9 | | | |
| | 11 | 5.4 | 5.5 | 5.6 | 5.7 | 5.9 | 6.0 | 6.1 | 6.3 | 6.4 | 6.5 | 6.7 | 6.8 | 6.9 | 7.0 | 7.2 | | | |
| | 12 | 5.6 | 5.8 | 5.9 | 6.0 | 6.1 | 6.3 | 6.4 | 6.5 | 6.7 | 6.8 | 6.9 | 7.1 | 7.2 | 7.3 | 7.4 | | | |
| | 13 | 5.9 | 6.0 | 6.2 | 6.3 | 6.4 | 6.5 | 6.7 | 6.8 | 6.9 | 7.1 | 7.2 | 7.3 | 7.5 | 7.6 | 7.7 | | | |
| | 14 | 6.2 | 6.3 | 6.4 | 6.6 | 6.7 | 6.8 | 6.9 | 7.1 | 7.2 | 7.3 | 7.5 | 7.6 | 7.7 | 7.9 | 8.0 | | | |
| | 15 | 6.4 | 6.6 | 6.7 | 6.8 | 7.0 | 7.1 | 7.2 | 7.3 | 7.5 | 7.6 | 7.7 | 7.9 | 8.0 | 8.1 | 8.3 | | | |
| | 16 | 6.7 | 6.8 | 7.0 | 7.1 | 7.2 | 7.4 | 7.5 | 7.6 | 7.7 | 7.9 | 8.0 | 8.1 | 8.3 | 8.4 | 8.5 | | | |
| | 17 | 7.0 | 7.1 | 7.2 | 7.4 | 7.5 | 7.6 | 7.8 | 7.9 | 8.0 | 8.1 | 8.3 | 8.4 | 8.5 | 8.7 | 8.8 | | | |
| | 18 | 7.2 | 7.4 | 7.5 | 7.6 | 7.8 | 7.9 | 8.0 | 8.2 | 8.3 | 8.4 | 8.5 | 8.7 | 8.8 | 8.9 | 9.1 | | | |
| | 19 | 7.5 | 7.6 | 7.8 | 7.9 | 8.0 | 8.2 | 8.3 | 8.4 | 8.6 | 8.7 | 8.8 | 8.9 | 9.1 | 9.2 | 9.3 | | | |
| | 20 | 7.8 | 7.9 | 8.0 | 8.2 | 8.3 | 8.4 | 8.6 | 8.7 | 8.8 | 9.0 | 9.1 | 9.2 | 9.3 | 9.5 | 9.6 | | | |
| | 21 | 8.1 | 8.2 | 8.3 | 8.4 | 8.6 | 8.7 | 8.8 | 9.0 | 9.1 | 9.2 | 9.4 | 9.5 | 9.6 | 9.7 | 9.9 | | | |
| | 22 | 8.3 | 8.5 | 8.6 | 8.7 | 8.8 | 9.0 | 9.1 | 9.2 | 9.4 | 9.5 | 9.6 | 9.8 | 9.9 | 10.0 | 10.1 | | | |
| | 23 | 8.6 | 8.7 | 8.9 | 9.0 | 9.1 | 9.2 | 9.4 | 9.5 | 9.6 | 9.8 | 9.9 | 10.0 | 10.2 | 10.3 | 10.4 | | | |
| | 24 | 8.9 | 9.0 | 9.1 | 9.3 | 9.4 | 9.5 | 9.6 | 9.8 | 9.9 | 10.0 | 10.2 | 10.3 | 10.4 | 10.6 | 10.7 | | | |
| | 25 | 9.1 | 9.3 | 9.4 | 9.5 | 9.7 | 9.8 | 9.9 | 10.0 | 10.2 | 10.3 | 10.4 | 10.6 | 10.7 | 10.8 | 11.0 | | | |
| | 26 | 9.4 | 9.5 | 9.7 | 9.8 | 9.9 | 10.1 | 10.2 | 10.3 | 10.4 | 10.6 | 10.7 | 10.8 | 11.0 | 11.1 | 11.2 | | | |
| | 27 | 9.7 | 9.8 | 9.9 | 10.1 | 10.2 | 10.3 | 10.5 | 10.6 | 10.7 | 10.8 | 11.0 | 11.1 | 11.2 | 11.4 | 11.5 | | | |
| | 28 | 9.9 | 10.1 | 10.2 | 10.3 | 10.5 | 10.6 | 10.7 | 10.9 | 11.0 | 11.1 | 11.2 | 11.4 | 11.5 | 11.6 | 11.8 | | | |
| | 29 | 10.2 | 10.3 | 10.5 | 10.6 | 10.7 | 10.9 | 11.0 | 11.1 | 11.3 | 11.4 | 11.5 | 11.6 | 11.8 | 11.9 | 12.0 | | | |
| | 30 | 10.5 | 10.6 | 10.7 | 10.9 | 11.0 | 11.1 | 11.3 | 11.4 | 11.5 | 11.7 | 11.8 | 11.9 | 12.0 | 12.2 | 12.3 | | | |
| | 31 | 10.8 | 10.9 | 11.0 | 11.1 | 11.3 | 11.4 | 11.5 | 11.7 | 11.8 | 11.9 | 12.1 | 12.2 | 12.3 | 12.4 | 12.6 | | | |
| | 32 | 11.0 | 11.2 | 11.3 | 11.4 | 11.5 | 11.7 | 11.8 | 11.9 | 12.1 | 12.2 | 12.3 | 12.5 | 12.6 | 12.7 | 12.8 | | | |
| | 33 | 11.3 | 11.4 | 11.6 | 11.7 | 11.8 | 11.9 | 12.1 | 12.2 | 12.3 | 12.5 | 12.6 | 12.7 | 12.9 | 13.0 | 13.1 | | | |
| | 34 | 11.6 | 11.7 | 11.8 | 12.0 | 12.1 | 12.2 | 12.3 | 12.5 | 12.6 | 12.7 | 12.9 | 13.0 | 13.1 | 13.3 | 13.4 | | | |
| | 35 | 11.8 | 12.0 | 12.1 | 12.2 | 12.4 | 12.5 | 12.6 | 12.7 | 12.9 | 13.0 | 13.1 | 13.3 | 13.4 | 13.5 | 13.7 | | | |
| | 36 | 12.1 | 12.2 | 12.4 | 12.5 | 12.6 | 12.8 | 12.9 | 13.0 | 13.1 | 13.3 | 13.4 | 13.5 | 13.7 | 13.8 | 13.9 | | | |
| | 37 | 12.4 | 12.5 | 12.6 | 12.8 | 12.9 | 13.0 | 13.2 | 13.3 | 13.4 | 13.5 | 13.7 | 13.8 | 13.9 | 14.1 | 14.2 | | | |
| | 38 | 12.6 | 12.8 | 12.9 | 13.0 | 13.2 | 13.3 | 13.4 | 13.6 | 13.7 | 13.8 | 13.9 | 14.1 | 14.2 | 14.3 | 14.5 | | | |
| | 39 | 12.9 | 13.0 | 13.2 | 13.3 | 13.4 | 13.6 | 13.7 | 13.8 | 14.0 | 14.1 | 14.2 | 14.3 | 14.5 | 14.6 | 14.7 | | | |

Word difficulty = percentage of words in the passage that are Sino-Vietnamese (of Chinese origin); the range is from 0% to 39%.
Sentence length = average number of words per sentence in the passage; the range is from 5 to 34.

results of the multiple regression analysis are presented in Table 2. The final equation is

$$RL = 1.74 + .226 + .129SL$$

The two predictors together account for almost 90% of the variance of the criterion. The revised readability formula for Vietnamese can be written:

$$RL = .27WD + .13SL + 1.74$$

The formula shows that the readability level (RL) of a piece of writing is accounted for by the predictors word difficulty (WD) and average sentence

length (SL).

The readability level (RL) is expressed in terms of grade level in the Vietnamese educational system, which is approximately equivalent to the U.S. system. The Vietnamese elementary-secondary system begins with grade 1 (age 6) and terminates at grade 12 (age 18).

Table 3 provides the readability level of Vietnamese prose when the word difficulty varies from 0 to 39 and average sentence length varies from 5 to 34.

Table 3 (continued)
Readability levels (grade levels, Vietnamese educational system)

| | Sentence length (average number of words) | | | | | | | | | | | | | | | |
|---|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | |
| Word difficulty (percentage of total words) | 0 | 4.3 | 4.5 | 4.6 | 4.7 | 4.9 | 5.0 | 5.1 | 5.2 | 5.4 | 5.5 | 5.6 | 5.8 | 5.9 | 6.0 | 6.2 |
| | 1 | 4.6 | 4.7 | 4.9 | 5.0 | 5.1 | 5.3 | 5.4 | 5.5 | 5.6 | 5.8 | 5.9 | 6.0 | 6.2 | 6.3 | 6.4 |
| | 2 | 4.9 | 5.0 | 5.1 | 5.3 | 5.4 | 5.5 | 5.7 | 5.8 | 5.9 | 6.0 | 6.2 | 6.3 | 6.4 | 6.6 | 6.7 |
| | 3 | 5.1 | 5.3 | 5.4 | 5.5 | 5.7 | 5.8 | 5.9 | 6.1 | 6.2 | 6.3 | 6.4 | 6.6 | 6.7 | 6.8 | 7.0 |
| | 4 | 5.4 | 5.5 | 5.7 | 5.8 | 5.9 | 6.1 | 6.2 | 6.3 | 6.5 | 6.6 | 6.7 | 6.8 | 7.0 | 7.1 | 7.2 |
| | 5 | 5.7 | 5.8 | 5.9 | 6.1 | 6.2 | 6.3 | 6.5 | 6.6 | 6.7 | 6.9 | 7.0 | 7.1 | 7.2 | 7.4 | 7.5 |
| | 6 | 6.0 | 6.1 | 6.2 | 6.3 | 6.5 | 6.6 | 6.7 | 6.9 | 7.0 | 7.1 | 7.3 | 7.4 | 7.5 | 7.6 | 7.8 |
| | 7 | 6.2 | 6.4 | 6.5 | 6.6 | 6.7 | 6.9 | 7.0 | 7.1 | 7.3 | 7.4 | 7.5 | 7.7 | 7.8 | 7.9 | 8.0 |
| | 8 | 6.5 | 6.6 | 6.8 | 6.9 | 7.0 | 7.1 | 7.3 | 7.4 | 7.5 | 7.7 | 7.8 | 7.9 | 8.1 | 8.2 | 8.3 |
| | 9 | 6.8 | 6.9 | 7.0 | 7.2 | 7.3 | 7.4 | 7.5 | 7.7 | 7.8 | 7.9 | 8.1 | 8.2 | 8.3 | 8.5 | 8.6 |
| | 10 | 7.0 | 7.2 | 7.3 | 7.4 | 7.6 | 7.7 | 7.8 | 7.9 | 8.1 | 8.2 | 8.3 | 8.5 | 8.6 | 8.7 | 8.9 |
| | 11 | 7.3 | 7.4 | 7.6 | 7.7 | 7.8 | 8.0 | 8.1 | 8.2 | 8.3 | 8.5 | 8.6 | 8.7 | 8.9 | 9.0 | 9.1 |
| | 12 | 7.6 | 7.7 | 7.8 | 8.0 | 8.1 | 8.2 | 8.4 | 8.5 | 8.6 | 8.7 | 8.9 | 9.0 | 9.1 | 9.3 | 9.4 |
| | 13 | 7.8 | 8.0 | 8.1 | 8.2 | 8.4 | 8.5 | 8.6 | 8.8 | 8.9 | 9.0 | 9.1 | 9.3 | 9.4 | 9.5 | 9.7 |
| | 14 | 8.1 | 8.2 | 8.4 | 8.5 | 8.6 | 8.8 | 8.9 | 9.0 | 9.2 | 9.3 | 9.4 | 9.5 | 9.7 | 9.8 | 9.9 |
| | 15 | 8.4 | 8.5 | 8.6 | 8.8 | 8.9 | 9.0 | 9.2 | 9.3 | 9.4 | 9.6 | 9.7 | 9.8 | 9.9 | 10.1 | 10.2 |
| | 16 | 8.7 | 8.8 | 8.9 | 9.0 | 9.2 | 9.3 | 9.4 | 9.6 | 9.7 | 9.8 | 10.0 | 10.1 | 10.2 | 10.3 | 10.5 |
| | 17 | 8.9 | 9.1 | 9.2 | 9.3 | 9.4 | 9.6 | 9.7 | 9.8 | 10.0 | 10.1 | 10.2 | 10.4 | 10.5 | 10.6 | 10.7 |
| | 18 | 9.2 | 9.3 | 9.5 | 9.6 | 9.7 | 9.8 | 10.0 | 10.1 | 10.2 | 10.4 | 10.5 | 10.6 | 10.8 | 10.9 | 11.0 |
| | 19 | 9.5 | 9.6 | 9.7 | 9.9 | 10.0 | 10.1 | 10.2 | 10.4 | 10.5 | 10.6 | 10.8 | 10.9 | 11.0 | 11.2 | 11.3 |
| | 20 | 9.7 | 9.9 | 10.0 | 10.1 | 10.3 | 10.4 | 10.5 | 10.6 | 10.8 | 10.9 | 11.0 | 11.2 | 11.3 | 11.4 | 11.6 |
| | 21 | 10.0 | 10.1 | 10.3 | 10.4 | 10.5 | 10.7 | 10.8 | 10.9 | 11.0 | 11.2 | 11.3 | 11.4 | 11.6 | 11.7 | 11.8 |
| | 22 | 10.3 | 10.4 | 10.5 | 10.7 | 10.8 | 10.9 | 11.1 | 11.2 | 11.3 | 11.4 | 11.6 | 11.7 | 11.8 | 12.0 | 12.1 |
| | 23 | 10.5 | 10.7 | 10.8 | 10.9 | 11.1 | 11.2 | 11.3 | 11.5 | 11.6 | 11.7 | 11.8 | 12.0 | 12.1 | 12.2 | 12.4 |
| | 24 | 10.8 | 10.9 | 11.1 | 11.2 | 11.3 | 11.5 | 11.6 | 11.7 | 11.9 | 12.0 | 12.1 | 12.2 | 12.4 | 12.5 | 12.6 |
| | 25 | 11.1 | 11.2 | 11.3 | 11.5 | 11.6 | 11.7 | 11.9 | 12.0 | 12.1 | 12.3 | 12.4 | 12.5 | 12.6 | 12.8 | 12.9 |
| | 26 | 11.4 | 11.5 | 11.6 | 11.7 | 11.9 | 12.0 | 12.1 | 12.3 | 12.4 | 12.5 | 12.7 | 12.8 | 12.9 | 13.0 | 13.2 |
| | 27 | 11.6 | 11.8 | 11.9 | 12.0 | 12.1 | 12.3 | 12.4 | 12.5 | 12.7 | 12.8 | 12.9 | 13.1 | 13.2 | 13.3 | 13.4 |
| | 28 | 11.9 | 12.0 | 12.2 | 12.3 | 12.4 | 12.5 | 12.7 | 12.8 | 12.9 | 13.1 | 13.2 | 13.3 | 13.5 | 13.6 | 13.7 |
| | 29 | 12.2 | 12.3 | 12.4 | 12.6 | 12.7 | 12.8 | 12.9 | 13.1 | 13.2 | 13.3 | 13.5 | 13.6 | 13.7 | 13.9 | 14.0 |
| | 30 | 12.4 | 12.6 | 12.7 | 12.8 | 13.0 | 13.1 | 13.2 | 13.3 | 13.5 | 13.6 | 13.7 | 13.9 | 14.0 | 14.1 | 14.3 |
| | 31 | 12.7 | 12.8 | 13.0 | 13.1 | 13.2 | 13.4 | 13.5 | 13.6 | 13.7 | 13.9 | 14.0 | 14.1 | 14.3 | 14.4 | 14.5 |
| | 32 | 13.0 | 13.1 | 13.2 | 13.4 | 13.5 | 13.6 | 13.8 | 13.9 | 14.0 | 14.1 | 14.3 | 14.4 | 14.5 | 14.7 | 14.8 |
| | 33 | 13.2 | 13.4 | 13.5 | 13.6 | 13.8 | 13.9 | 14.0 | 14.2 | 14.3 | 14.4 | 14.5 | 14.7 | 14.8 | 14.9 | 15.1 |
| | 34 | 13.5 | 13.6 | 13.8 | 13.9 | 14.0 | 14.2 | 14.3 | 14.4 | 14.6 | 14.7 | 14.8 | 14.9 | 15.1 | 15.2 | 15.3 |
| | 35 | 13.8 | 13.9 | 14.0 | 14.2 | 14.3 | 14.4 | 14.6 | 14.7 | 14.8 | 15.0 | 15.1 | 15.2 | 15.3 | 15.5 | 15.6 |
| | 36 | 14.1 | 14.2 | 14.3 | 14.4 | 14.6 | 14.7 | 14.8 | 15.0 | 15.1 | 15.2 | 15.4 | 15.5 | 15.6 | 15.7 | 15.9 |
| | 37 | 14.3 | 14.5 | 14.6 | 14.7 | 14.8 | 15.0 | 15.1 | 15.2 | 15.4 | 15.5 | 15.6 | 15.8 | 15.9 | 16.0 | 16.1 |
| | 38 | 14.6 | 14.7 | 14.9 | 15.0 | 15.1 | 15.2 | 15.4 | 15.5 | 15.6 | 15.8 | 15.9 | 16.0 | 16.2 | 16.3 | 16.4 |
| 39 | 14.9 | 15.0 | 15.1 | 15.3 | 15.4 | 15.5 | 15.6 | 15.8 | 15.9 | 16.0 | 16.2 | 16.3 | 16.4 | 16.6 | 16.7 | |

Validity and reliability

The validity of a readability formula may be suggested by a high correlation between the score yielded by the formula and those obtained from a comprehension test or from the judgments of experts. The second method, correlation between the formula and expert judgments, has been applied in this study.

Thirty Vietnamese passages of about 100 words each were selected randomly from books and magazines. Six former teachers of Vietnamese

judged the readability level of each passage. All six teachers have had more than 10 years of teaching experience, and all hold bachelor's degrees specializing in the teaching of Vietnamese.

Statistics related to the teachers' rating scores and the score yielded by the new readability formula are presented in Table 4. The correlation between the criterion scores and the formula is high ($r = .92$), ascertaining the validity of the formula based on the correlational analysis.

Table 4
Statistics related to the teachers' rating scores and
scores yielded by the new readability formula (N = 30)

| Variables | Mean | s.d. | r |
|---|--------|--------|--------|
| Teachers' ratings | 7.3380 | 2.4163 | |
| New formula's ratings | 6.9077 | 3.1812 | |
| Teacher's ratings, new formula's ratings | | | .9220* |

* $p < .01$

Another sample of 31 passages with two versions (one in English and one in Vietnamese) was selected randomly from bilingual materials (originally written in English and translated into Vietnamese, or vice versa) for a second study. The readability levels of the English versions of these passages were assessed using the Flesch formula (Flesch, 1948). The readability levels of the Vietnamese versions were measured using the revised formula.

The correlation between the Flesch formula and the revised formula is $-.9252$. Since the translations were carefully executed by Vietnamese experts, it is assumed that the two versions of these 31 passages are equivalent in terms of readability. Thus, the high correlations between the Flesch formula and the revised formula may be viewed as supporting the validity of the latter.

The reliability of a readability formula focuses on the extent to which different people who use the same formula to measure the same passages come up with the same results. To determine the reliability of the revised formula, the following method was adopted. Twelve passages of about 100 words each were selected randomly from Vietnamese books and magazines. They were sent to four Vietnamese teachers who used the revised readability formula to determine

their readability levels. Two of these teachers (T1 and T2) taught at the secondary level and specialized in Vietnamese language arts. The third teacher (T3) taught English as a second language. The fourth (T4) was an elementary school teacher with specialization in mathematics and science.

The correlations among the scores of the four teachers on the 12 passages are presented in Table 5. The average correlation is $.9711$, and the reliability coefficient ($.94$) is relatively high.

How to use the formula

A basic knowledge of Vietnamese will suffice for those who use the formula to compute the difficulty of a passage. Select passages of about 100 words each, about three selections per text. Count the words in each passage. Count the sentences per passage. Compute the average sentence length (SL) of each passage by dividing the total number of words by the total number of sentences.

$$SL = \frac{\text{total number of words}}{\text{total number of sentences}}$$

Compute the percentage of compound Sino-Vietnamese words (WD) in each passage, using the formula

$$WD = \frac{\text{number of compound Sino-Vietnamese words}}{\text{total number of words}} \times 100$$

Table 5
Correlations among teachers' scores (N = 12)

| | Teacher 1 | Teacher 2 | Teacher 3 | Teacher 4 |
|-----------|-----------|-----------|-----------|-----------|
| Teacher 1 | 1.0000 | | | |
| Teacher 2 | .9991 | 1.0000 | | |
| Teacher 3 | .9731 | .9783 | 1.0000 | |
| Teacher 4 | .9623 | .9644 | .9497 | 1.0000 |

A distinguishing characteristic of compound Sino-Vietnamese words relates to ordering. The adjective or adverb precedes the noun or verb in a compound Sino-Vietnamese word, for example, *thanh-thiên* 'blue sky,' or *trúc-thang* 'vertically, to go up/ascend,' literally, 'helicopter.'

A compound Vietnamese word, in contrast, appears in the opposite order, for example, *trời-xanh* 'sky blue,' or *lên-thăng* 'to go up/ascend, vertically.' The Sino-Vietnamese word for airplane is *phi-cô* 'flying engine'; the Vietnamese compound equivalent is *máy-bay* 'engine-flying.'

When SL and WD are obtained, use these values to find the readability level of the passage using Table 3. When using Table 3, if the value of SL is less than 5, consider it as 5; if the

value of SL is larger than 34, consider it as 34. If the value of WD is larger than 39, consider it as 39. If the value of RL is 13 or above, consider it as college level.

The revised formula is easy to use and yields valid and reliable results.

References

- Brinton, James, and Wayne Danielson. "A Factor Analysis of Language Elements Affecting Readability." *Journalism Quarterly*, vol. 35 (Fall 1958), pp. 420-26.
- Flesch, Rudolf. "A New Readability Yardstick." *Journal of Applied Psychology*, vol. 32 (June 1948), pp. 221-33.
- Nguyen, Liem, and Alan Henkin. *Fundamental Vietnamese Vocabulary*. Iowa City, Iowa: National Center for Materials and Curriculum Development, The University of Iowa, 1981a.
- Nguyen, Liem, and Alan Henkin. "A Readability Formula for Vietnamese." *Journal of Reading*, vol. 26 (December 1982), pp. 243-51.
- Nguyen, Liem, and Alan Henkin. *Vietnamese Refugee Students: Legacies of an Educational Past*. Iowa City, Iowa: National Center for Materials and Curriculum Development, The University of Iowa, 1981b.
- Stolurow, Lawrence, and Robert J. Newman. "A Fractional Analysis of Objective Features of Printed Language Presumably Related to Reading Difficulty." *Journal of Educational Research*, vol. 52, no. 7 (1959), pp. 243-51.

"Yellow Pages" for principals

During the 1982-83 and 1984-85 school years, the staff of the Instructional Management Program at the Far West Laboratory promoted an information exchange by telephoning principals. The interviews lasted 1 to 2 hours and centered around a description of programs and practices principals used in dealing with some common instructional management problems.

The results of the information exchanges were two handbooks of administrative techniques, in the form of catalogs, one for elementary principals and one for secondary principals. The catalogs include topics such as inservice and training for new teachers; monitoring student achievement in reading; strategies for improving staff morale; and using computers for remedial reading. Catalog users can read about the advantages and disadvantages of the practices listed. An index enables readers to locate the names and telephone numbers of colleagues who use the practices and call them to discuss the actual implementation.

For information, write Far West Laboratory for Educational Research and Development, 1855 Folsom St., San Francisco, CA 94103, USA. Phone (415) 565-3000.