Do Syllable Count and Word Frequency Differ Significantly in Easy and Difficult Reading Comprehension Items?

Kyle Perkins
Florida International University, USA

Abstract: This study investigated the effect of the number of syllables and the word frequency of the words in the reading passages, the question stems, and the answer options of easy and difficult reading comprehension items. Significant differences were found for the easy and difficult items.

The purpose of this study was to determine whether the number of syllables and the word frequency of the words in the reading passages (texts), the question stems, and the answer options in a reading comprehension test were significantly different for easy and difficult reading comprehension items for English language learners (ELLs). This research investigates two factors that may influence an ELL's reading processing that occurs in working memory. "Working memory refers to the information that is activated, or given mental stimulation, for immediate storage and processing" (Grabe & Stoller, 2002, p. 18).

Humans have a limited capacity for language and speech processing. "Humans are designed to process material one element at a time (think of the linear aspect of language, for example, we hear and process one sound, one word, at a time)" (Sobel, 2001, p. 48). Our limited information processing capacity may be due to working memory. The working memory can process only five to nine pieces of information at any given moment (Miller, 1956). More recent research indicates that the number may be nearer to three or four (Feldon, 2010). The working memory organizes and processes incoming information and interacts with knowledge in the long-term memory. Given the working memory's maximum capacity of nine pieces of information (perhaps, an over estimation), it is limited to processing no more than two or three relationships at once (Novak & Canas, 2008). Crain and Shankweiler (1988) showed that sentence length is a surrogate measure of structural complexity demands on the working memory. A reader's limited capacity for language processing suggests that the more words a reader must process during a timed test of reading comprehension, the more a reader's cognitive capacity might constrain comprehension, and hence some test items might be difficult as a result.

For over three decades, the top-down and the bottom-up theories of reading have dominated research and the didactic literature. According to the bottom-up proponents, "reading is about processing letters and words" (Pressley, 1998, p. 52). Readers are presumed to process letters and words systematically and thoroughly (Gough, 1972). The more letters and words to be processed, the more time a reader must spend on the reading task.

In marked contrast to the bottom-up proponents, the top-down theorists believe that "based on world knowledge, people have hypotheses about what the text is going to say, and this prior knowledge goes far in explaining comprehension" (Pressley, 1998, p. 53). A single text can have different meanings for different readers because meaning is the end product of three inputs: the author's words and text, and the reader's prior knowledge (Beach & Hynds, 1991). Grabe and Stoller (2002), Herrera, Perez, and Escamilla (2010), and Lems, Miller, and Soro (2010) provide further explication of the top-down, bottom-up, and integrated models of reading comprehension and their relationship to teaching reading to ELLs.

Perkins, K. (2001). Do syllable count and word frequency differ significantly in easy and difficult reading comprehension items? In M. S. Plakhotnik, S. M. Nielsen, & D. M. Pane (Eds.), *Proceedings of the Tenth Annual College of Education & GSN Research Conference* (pp. 177-182). Miami: Florida International University. http://coeweb.fiu.edu/research_conference/

Many top-down theorists eschew a reader's systematic processing of letters and words, claiming that readers only engage in bottom-up processing, if and when they experience difficulty in meaning-making with the text. The results from the eye-movement research paradigm suggest otherwise. For example, Perfetti (1985) and Stanovich (1980) reported that even skilled readers use orthographic information to identify words. Readers sample nearly seventy-five percent of the content words and approximately fifty percent of the function words in a text. Treiman (2001) claimed that even skilled readers fixate on a majority of a text's words because we humans have a fairly small span of useful vision—an anatomical feature that results in a limited amount of data for processing by the working memory. The research reviewed here suggests that skilled and less-skilled readers sample a substantial number of words in a text, a question stem, and a list of answer options. A larger number of words then could well constrain comprehension and result in difficult items.

The frequency with which the words in a text, a question stem, and a list of answer options occur in written and oral communication in the target language could also have an effect on the reading processing that goes on in an ELL's working memory during the reading comprehension process. Stenner, Burdick, Sanford, and Burdick (2006) proposed the concept of exposure theory to explain the effect that word frequency has on the development of a reader's receptive vocabulary. The basic idea is that the more frequently a word appears in written and oral communication, the higher the probability that such a word will become a member of a reader's receptive vocabulary.

By extension, it would seem to follow that if readers have an extensive, receptive vocabulary, and that if the words they process in a reading comprehension test are frequent and familiar, then the readers should have faster access to these words, resulting in more efficient processing of the incoming data. Words that readers don't know because these words have been encountered less frequently and have not been added to the receptive vocabulary could prove problematic and tax the comprehension process. For many generations, reading teachers have encouraged their students to guess a word's meaning from its context, but, as Treiman (2001) has noted, many words are only minimally predictable from context, if at all.

Based on the preceding review of the literature, I developed four research hypotheses: (a) The number of syllables in the reading passages (texts), the question stems, and the answer options would be significantly different for the easy and the difficult reading comprehension items. For this research, the number of syllables to be processed during reading comprehension is assumed to be a proxy measure of a reader's information processing load. (b) The higher number of syllables should be associated with the difficult items. (c) The word frequency of the words in the reading passages (texts), the question stems, and the answer options would be significantly different for the easy and the difficult items. (d) The higher frequency words should be associated with the easy items.

The four research hypotheses derive from the following purported if-then conditions. If the number of syllables to be processed and the number of low frequency words not found in a reader's receptive vocabulary pose an overload to the reader's working memory, then the working memory may lose information to be processed, and reading comprehension fails, or is severely compromised.

Method

Data

The data analyzed for the present study were the scored item responses in the reading comprehension section of a Test of English as a Foreign Language (TOEFL) test form, which

was used in an institutional administration at a rural university in the Midwestern U.S. The 202 examinees averaged 456.99 (s=59.51) on the overall TOEFL. Twenty-five native languages were represented in the subject pool. The examinees comprised a cohort of international students who were enrolled in an intensive English language institute. The examinees planned to enroll in American universities after having attained the requisite acquired English proficiency for admission to full-time university study. This university setting was chosen because over 200 ELLs representing over 20 native languages were available for data collection.

The reading comprehension section contained five reading passages, which ranged from 205 to 339 words in length, and the number of questions per passage ranged from five to eight.

Analyses

I converted the raw score summaries to their natural logarithms in order to produce perfect interval linear measures. The natural logarithms of the person and item success-to-failure ratios are necessary to represent the relative distances between raw scores. For example, the measure distance between the scores of 88 percent correct and 98 percent correct is 4.75 times greater than the distance between the scores of 45 percent correct and 55 percent correct.

For this study, an easy item was defined as having a positive natural log and appearing above the midpoint (0.0) of the person ability and item difficulty interval scale. A difficult item was defined as having a negative natural log and appearing below the midpoint of the person ability and item difficulty interval scale.

The frequency of each word in the reading passages (the texts), the question stems, and the answer options was obtained from the Carroll, Davies, and Richman (1971) frequency list. The Carroll et al. study determined the frequency with which words occurred (more appropriately, re-occurred) in a five million-word corpus of running text. For each of the three sections of the test (reading passages, question stems, and the answer options), the word frequencies were summed, and the natural log of that sum was entered into the analyses.

The Mann-Whitney *U* test was utilized to analyze the data because neither the homogeneity of the variance nor the normality of distribution could be guaranteed for the parametric family of statistical procedures. A large number of tie scores was encountered while calculating the Mann-Whitney *U* statistics; therefore, the normal approximation with tie correction was employed. The calculation of the statistic with tie correction produces a *z* score.

Results

Table 1 presents the item difficulties and the person abilities for the 29 items and the 202 students who sat for the exam. Table 1 shows that there were ten easy items and six difficult items. The easy items were questions 31, 32, 35, 38, 41, 42, 43, 44, 46, and 49. The difficult items were questions 39, 50, 52, 53, 54, and 59. There were three easy items associated with the first reading passage and no difficult items; the second passage, four easy items and one difficult item; the third passage, two easy items and no difficult items; the fourth passage, one easy item and four difficult items; and the fifth passage, no easy items and two difficult items.

Table 2 presents the results of the statistical tests for the differences in the ranks of the easy and the difficult items. There were significant differences in the ranks of the easy and the difficult items for the number of syllables in the texts, the question stems, and the answer options. The question stems and the answer options in the difficult items contained more syllables than their counterparts in the easy items, as was predicted. However, the texts associated with the easy items contained more syllables than their counterparts in the difficult items, which was exactly the opposite of what was predicted in the second research hypothesis.

There were significant differences in the ranks of the easy and the difficult items for the word frequencies of the words in the texts and the word frequencies of the words in the answer options, as was predicted. However, the word frequencies of the texts and the answer options were higher for the difficult items, which was the opposite of what was predicted in the second research hypothesis. No significant difference was found in the word frequencies in the question stems for the easy and difficult items.

Discussion and Conclusion

The results may lend support to the idea that a syllable count is a measure of processing load for ELL students sitting for the TOEFL reading comprehension test and that significantly more syllables in the question stems and in the answer options may make such items more difficult. The answer options for two questions illustrate the point: the options for one question were four single words, and another question had a complete sentence for each of the four answer options, a total of 70 words.

The fact that the texts associated with the difficult items may be a tempest in a teapot, i.e., not significant in the greater scheme of things for two reasons. First, there was a difference of only 134 words between the longest reading passage and the shortest reading passage. Second, there may be other factors that determine easy and difficult items.

Word frequency may not have the same explanatory power for non-native speakers sitting for a foreign language proficiency test as it does for persons who have native proficiency in the language being tested. Stenner et al.'s (2006) exposure theory is based on the idea that the more occasions a person is exposed to a word, the greater is the likelihood that that word will become part of that person's receptive vocabulary. But mere exposure to a word does not guarantee its meaningful learning, i.e., that the word or concept to which it refers becomes assimilated into a person's existing cognitive structures. Cognitive scientists, e.g., Ausubel (1963, 1968), Ausubel et al. (1978), and Novak and Canas (2008), define meaningful learning as the successful assimilation of new concepts and propositions into existing cognitive structures.

Meaningful learning requires three conditions: (a) the material to be learned must be conceptually clear and presented with language and examples relatable to the learner's prior knowledge; (b) the learner must possess relevant prior knowledge; and (c) the learner must choose to learn meaningfully…Individuals may vary in the quantity and quality of the relevant knowledge they possess, and in the strength of their motivation to seek ways to incorporate new knowledge into relevant knowledge they already possess (Novak & Canas, 2008, pp. 2-3).
It may indeed be the case that the students who provided the data for this research had not yet meaningfully learned enough words for word frequency to be a reliable and valid discriminator of easy and difficult reading comprehension items.

The Mann-Whitney *U* test indicated no significant differences in the ranks of the easy and the difficult items for the word frequencies of the words in the question stems. An examination of the question stems revealed that they comprised sentence completions, direct WH-questions, and embedded WH-questions. In other words, the sentence types (sentence structures), constituent structures, and the slot-and-filler sentence patterns were similar for both the easy and the difficult questions.

The results present one clear finding: the question stems and the answer options in the difficult items contain significantly more syllables than their counterparts in the easy items. These extra syllables constituted an increased load in the test-takers' information processing, and this is one factor (but not a complete explanation) for the reading items' difficulty. The remaining results are a mixed grill. The easy items associated with texts that have significantly

more syllables than the texts that are associated with the difficult items. Although the difference was significant, it may be not be meaningful because there were no tremendous differences in length for any of the passages. The differences in word frequencies may also be of little relevance because the test-takers may not have meaningfully learned enough words, at the time they sat for the test, for word frequency to be a valid discriminator between easy and difficult items.

By selecting syllable counts and word frequency as the foci for this research, I did not assume that these two variables were the sole determinants of item difficulty. Reading is a multidimensional construct. "Competent reading is an integrative and functional act; that is, it requires successfully combining (integrating) a number of skills for the purpose of accomplishing concrete goals (functions)" (Goldenberg, in press, 2010, pp. 21-22).

References

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York, NY: Grune and Stratton.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York, NY: Holt, Rinehart and Winston.

Ausubel, D. P., Novak, J. D., & Hanesian. H. (1978). *Educational psychology: A cognitive view.* (2nd ed.). New York, NY: Holt, Rinehart and Winston.

Beach, R., & Hynds, S. (1991). Research on response to literature. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research, Vol. II* (pp. 453-489). White Plains, NY: Longman.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The word frequency book.* Boston, MA: Houghton Mifflin.

Crain, S., & Shankweiler, D. (1988). Syntactic complexity and reading acquisition. In A. Davidson & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 167-192). Hillsdale, NJ: Lawrence Erlbaum.

Feldon, D. F. (2010). Why magic bullets don't work. *Change, 42*(2), 15-21.

Goldenberg, C. (in press). Reading instruction for English language learners. In M. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. IV). Newark, NJ: International Reading Association.

Gough, P. B. (1972). One second of reading. In J. F. Kavanagh, & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 351-358). Cambridge, MA: MIT Press.

Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading.* London, UK: Longman.

Herrera, S. G., Perez, D. R., & Escamilla, K. (2010). *Teaching reading to English language learners: Differentiated literacies.* Boston, MA: Allyn & Bacon.

Lems, K., Miller, L. D., & Soro, T. M. (2010). *Teaching reading to English language learners: Insights from linguistics.* New York, NY: The Guilford Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Novak, J. D., & Canas, A. J. (2008). The theory underlying concept maps and how to construct and use them. Retrieved from http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptsMaps.htm

Perfetti, C. A. (1985). Reading ability. In R. J. Sternberg (Ed.), *Human abilities: An information-processing approach* (pp. 59-81). New York, NY: W. H. Freeman.

Pressley, M. (1998). *Reading instruction that works: The case for balanced teaching.* New York, NY: The Guilford Press.

Sobel, C. P. (2001). *The cognitive sciences: An interdisciplinary approach.* Mountain View, CA: Mayfield.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32-71.

Stenner, A. J., Burdick, H., Sanford, E. S., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement, 7*(3), 307-322.

Treiman, R. (2001). Reading. In M. Aronoff, & J. Rees-Miller (Eds.), *Blackwell handbook of linguistics* (pp. 664-672). Oxford, UK: Blackwell.

Table 1

Item Difficulty and Person Ability Natural Log Scale

The number of persons at a particular person ability level	Natural log	Items at a particular item difficulty level-test item numbers
n = 1	3.5	
	3.0	
	2.5	
n = 8	2.0	46
n = 17	1.5	43, 49
n = 21	1.0	31, 32, 42
n = 36	0.5	35, 38, 41, 44
n = 89	0.0	33, 34, 36, 37, 40, 45, 47, 48, 51, 55, 56, 57, 58
	-0.5	50, 52, 53
n = 21	-1.0	39, 54, 59
n = 4	-1.5	
n = 5	-2.0	
	-2.5	

Table 2

Differences in the Ranks of Difficult and Easy Items

Variable	Test statistic	Mean of the easy items	Mean of the difficult items
Number of syllables in the text	z = -1.65 p < .05	428.10	383.50
Number of syllables in the question stems	z = -1.04 p < .05	19.80	21.83
Number of syllables in the answer options	z = -2.07 p < .05	16.80	40.33
Word frequencies of the texts	z = -2.29 p < .05	16.55	16.77
Word frequencies of the question stems	U = 30 n.s.	13.68	13.88
Word frequencies of the answer options	z = -1.19 p < .05	9.17	11.50