Text Variables as Predictors of Reading Item Difficulty

Northern Arizona University

Sumi Han

Abstract

Automated text complexity measures such as the Lexile measure and Coh-Metrix have helped teachers determine passages appropriate for their students. From a perspective of testing, the study examined if the Lexile measure and 11 text complexity variables can predict reading item difficulty of PIE (Program in Intensive English at Northern Arizona University) placement reading tests. During reading test development, the Lexile measure is used as a criterion of text difficulty for selecting level-appropriate reading passages. It turned out that the Lexile measure predicts 34 % of the item difficulty (i.e., equated deltas) of two placement reading tests. Connectives (e.g., transition words) and the ratio of academic words were also significant predictors of reading item difficulty. It is expected that the results and findings of the present study can help teachers find or adjust reading passages for testing purposes in the PIE.

Background

Automated text complexity measures have helped teachers determine passages appropriate for their students in teaching and testing (Fisher, Frey, & Lapp, 2012). There are free online text complexity measures such as the Lexile measure (MetaMetrics) and Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004), and in-house text complexity software programs such as SourceRater (Educational Testing Service) and Pearson Reading Maturity Metric (Landauer, 2011). Each of the text complexity measures is based on different quantitative (e.g., word frequency, sentence length) and qualitative features (e.g., text cohesion) of text complexity.

The Lexile measure is currently used in the PIE (Program in Intensive English) reading test development, but it has not been empirically examined if it appropriately predicts item difficulty of placement reading tests. It is also assumed that two variables of word frequency and sentence length underlying the Lexile measure would not explain the whole picture of examinees' reading proficiency. To fill this gap, potential predictors of item difficulty were pooled from variables used in other automated text complexity measures and examined for their explanatory power of reading item difficulty.

Research Questions

The following two research questions were raised for the current study:

1. Does the Lexile measure predict reading item difficulty of the PIE placement reading tests?
2. What other text variables predict reading item difficulty of the PIE placement reading tests?

Methods

Participants

A total of 332 participants took the Fall 2010, 2011, and 2012 PIE Placement tests. Most of them were male ESL students, who just came from two countries, Saudi Arabia and China at the time

of test administration. Their English proficiency varied from beginners to advanced non-native learners of English.

Reading Sections of the Fall Placement Tests

The PIE placement tests assessed the examinees' academic English language proficiency and the results were used to place them into appropriate levels of class in the PIE. The reading sections of the three Fall Placement tests had 13 reading passages with 110 items. Different reading passages in each reading section represented different levels of reading proficiency, and a total of six objectives were targeted in the reading test: vocabulary, details, and main ideas for basic comprehension and text organization, inference, and purpose/attitude for advanced comprehension. For the basic text complexity information, estimates of word count, the Lexile measure, and the Flesch-Kincaid Grade Level (FKGL) were given per passage.

Dependent and Independent Variables

The equated delta per item in the three tests was used as the dependent variable. Equated delta has a strong negative relation with item difficulty (i.e., percent correct per item), meaning the higher the equated delta of an item is, the more difficulty the item is. The independent variables are divided into two categories according to the research questions. For the first research question, there were three independent variables: the Lexile measure and its underlying variables, log mean sentence length, and mean of log word frequency. For the second research question, there were a total of 11 variables of text complexity based on text complexity measures (e.g., Fresch-Kincaid, Coh-Metrix, SourceRater, and Pearson Reading Maturity Metric) and categorized into three groups: Lexical, Syntactic, and Semantic Category. A detailed description of each of the variables is given in Table 1. Each of the variables was automatically calculated per passage, using Tob Cobb's Vocab Profile and Coh-Metrix.

Analyses

The three reading sections had one common reading passage *Bioluminescence,* which was used as a linking passage. In the 2012 Placement test, the overall reading section and the reading passage *Bioluminescence* had better item discriminability power and normal distribution compared to the passage in the other tests. Thus, it was used as a reference group for delta equating. Delta equating was performed on the item difficulty values of the same passage used in the three reading sections in order to eliminate the effect of group ability on item difficulty (Livingston, 2013). First, the distribution of item difficulty (i.e., percent correct per item) was transformed to be normal (Mean = 13, SD = 4), which generated deltas of all the items in each reading section. Second, delta equating was done on the items between the Fall 2012 linking passage (as a reference group) and each of the Fall 2010 and Fall 2011 linking passage (as an observed group). Third, the contrast between the averaged delta and equated delta for the Fall 2010 linking passage and for the Fall 2011 linking passage was applied to the delta values of the other passages in the same reading section. In the end, correlation and hierarchical and stepwise multiple regression analyses were performed on the reading test data.

Results

Tables 2 to 5 list results from correlation and multiple regression analyses on the reading test data. Regarding the first research question, Table 2 showed that there exit moderate correlations between the equated deltas and each of the independent variables, the Lexile Measure (Lexile), LMSL (log mean sentence length), and LWF (mean of log word frequency). Lexile is highly correlated with each of its underlying variables, LMSL and LWF. Table 3 demonstrated that 34 % of the variance in the equated deltas explained only by the Lexile measure.

Table 1

Independent Variables of Text Complexity Used in Study

Lexical Category		**Definition**
Lex1	Word length	Syllables per word
Lex2	Word frequency	Total word counts
Lex3	Academic Words*	Academic word counts/Total word counts
Lex4	Off-list word frequency*	Word counts on the off-list/Total word counts
Lex5	Type/Token Ratio	# of unique words/Total word counts (for content words)
Syntactic Category		**Definition**
Syn1	Sentence Length	# of total word counts/# of sentences
Syn2	NP density	# of modifiers/# of noun phrases
Syn3	Passive	# of passive voice structures/Total word counts
Semantic Category		**Definition**
Sem1	Co-reference cohesion	% of all sentence pairs sharing a common noun, pronoun or noun phrase
Sem2	LSA	Text cohesion based on Lexical Semantic Analysis
Sem3	Connectives	# of connectives (e.g., transition words)/Total word counts

Note. * = Tob Cobb's Vocab Profile; others from Coh-Metrix version 3.0

Table 2

Correlation Coefficients Among Equated Delta, Lexile, LMSL, and LWF

	Lexile	LMSL	LWF
Equated Delta	.58	.56	-.52
Lexile		.98	-.86
LMSL			-.74

Note. All significant at p=.000

Table 3

Results of Hierarchical Multiple Regression

Model	Total R^2	ΔR^2
1	.340***	.340***

Note. Predictors: Lexile, ***: p = .000

Regarding the second research question, Tables 4 and 5 showed significant predictors of item difficulty (i.e., equated delta) other than the word frequency (Lex2) and mean sentence length (Syn1). Especially, Academic words (Lex3) and Connectives (Sem3) were closely related with the equated deltas. The more use of academic words and less use of connectives (e.g., in other words, however) in the passage, the more difficult items are.

Table 4

Correlation Coefficients Among Equated Delta and Each of 11 Independent Variables

	Lex1	Lex2	Lex3	Lex4	Lex5	Syn1	Syn2	Syn3	Sem1	Sem2	Sem3
Equated Delta	.52	.46	.34	.23	.23	.40	*.04*	.37	-.19	*.15*	-.53

Note. All significant (except for the italicized values) at p = .05 (1-tailed)

Table 5 listed results from stepwise multiple regression with the equated deltas and the 11 variables. LSA and Connectives (Sem3 and Sem2) are potential predictors of item difficulty: They, with lexical variables (Lex1, Lex2, Lex3), explain 48% of variance in item difficulty. The ratio of academic words (Lex3) has 10% unique contribution to explaining the variance of item difficulty.

Table 5

Results of Stepwise Multiple Regression

Model	Total R^2	ΔR^2	Partial Coefficients (*B*)				
			Sem3	Lex3	Lex2	Lex1	Sem2
1	.277	.277	-15.52				
2	.385	.108	-15.22	38.32			
3	.411	.026	-8.05	44.97	.004		
4	.438	.027	6.77	24.61	.008	7.44	
5	.478	.040	18.41	*15.39*	.012	12.12	6.29

Note. All significant (except for the italicized value) at p = .05

Relevance to PIE and Second Language Learning

The results and findings of the present study can help teachers find or adjust reading passages for testing purposes. During PIE reading test development, the Lexile measure can be continuously used as a text complexity measure, predicting reading item difficulty. Cohesion and academic words can be also used as potential predictors of reading item difficulty during test development.

References

Fisher, D., Frey, N., & Lapp, D. (2012). *Text complexity: Raising rigor in reading*. Newark, DE: International Reading Association.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*, 193-202.

Landauer, T. K. (2011). *Pearson's text complexity measure*. Pearson. Retrieved on April 25, 2013 from http://www.readingmaturity.com.

Livingston, S. A. (2013). *Delta equating.* Educational Testing Service.