
Poster Presentation

Using Text Analysis Measures to Aid in EFL Content Textbook Selection

Dax Thomas

Meiji Gakuin University

When faced with the task of choosing textbooks for their English as a Foreign Language (EFL) content classes, many teachers approach the decision based on an intuitive overall feel for the appropriateness and difficulty level of the text. While this is often perfectly satisfactory, a quantitative approach towards determining textbook appropriateness may help bolster teachers' confidence in their textbook choices.

This paper reports on the use of a combination of text analysis measures to aid in the selection of a content textbook for an intermediate level themed EFL class at a Japanese university. Three potential content textbooks were analysed using three different sets of text analysis measures: readability, lexical diversity, and vocabulary coverage. Based on these measures, the texts could be easily ranked according to their appropriateness for the target level of the themed EFL class.

Textbook Context

All students in the Centre for Liberal Arts department for which this course was conducted are required to take at least one content-themed EFL class over the course of their four-year program. This is usually done in the second semester of their freshman year, and instructors offer courses based on their own personal academic background. The course referred to in this paper is "Introduction to Archaeology" and is offered as an intermediate level "English Seminar" class.

The Ideal Textbook

For this course, a textbook with five specific characteristics, or criteria, was desired:

- a. It should be written in natural English (i.e., targeted at native speakers).
- b. It should be accessible to intermediate level university EFL students (i.e., not too difficult).
- c. It should have as much vocabulary recycling as possible (i.e., low lexical diversity).
- d. It should have a fair coverage of General Service List (GSL) and Academic Word List (AWL) vocabulary.
- e. It should have good coverage of important content-related vocabulary (i.e., archaeology-related words).

Three texts were chosen as possible textbooks for use in this class: *Archaeology: Discovering the Past* by John Orna-Ornstein (2002), *Archaeology (Kingfisher Knowledge)* by Trevor Barnes (2004), *Archeology (Eyewitness Books)* by Jane McIntosh (2000). These will later be referred to as the Orna text, Barnes text, and McIntosh text.

Each of these texts are targeted at native English speaking younger learners (middle school to early high school) and consist of between 40 and 60 pages of text with a high number of photographs and illustrations. Each of the books was chosen because its length and general subject matter appeared appropriate to the course and level of the students.

Text Analysis Measures

The measures utilised in this study were readability, lexical diversity, and vocabulary coverage.

Readability

The three readability indices chosen for this study were the Flesch-Kincaid Readability Index (FKRI; Flesch, 1948), the Flesch-Kincaid Grade Level (FKGL; Kincaid, Fishburne, Rogers, & Chissom, 1975), and the Simple Measure of Gobbledygook Index (SMOG; McLaughlin, 1969). All three were calculated using the software Readability (Silverstein, 2009). These are well-known measures

and are generally considered appropriate for longer secondary and tertiary level texts. The calculation for the FKRI and the FKGL are both based on average sentence length and average number of syllables per word. FKRI outputs a score ranging between 0 and 100 (a higher score indicating an easier text). The score output by FKGL reflects what U.S. grade level that text would be appropriate for. The calculation for the SMOG is based on the square root of the polysyllabic word count of sections of the text and, like the FKGL, the score represents a U.S. grade level.

Lexical Diversity

Two lexical diversity measures, the Type-Token Reference Curve Index (TTRCI; Thomas, 2016) and the Type-Token Area Index (TTAI; Thomas, 2016), were applied to the texts. These are two relatively new measures of the type-token relationship. In the TTRCI, the type-count of a text is compared against a point on a reference (normative) curve at the same token count. The difference between these type-counts is used in calculating the index. TTAI takes the difference between the area under the type-token curve of a text and the area under a type-token reference curve of the same token length, to form the index. This measure attempts to take into account shifts in lexical diversity throughout the text¹.

Vocabulary Coverage

Vocabulary coverage of the General Service List (GSL; first 1,000 and second 1,000 words; West, 1953) and the Academic Word List (AWL; 570 words; Coxhead, 2000) was analysed using the software AntWordProfiler (Anthony, 2014). In order to analyse the vocabulary related to the content theme of archaeology, frequency band word lists were created from a commonly used introductory archaeology textbook targeting first-year native English speaking university students. The textbook chosen for this was *The Archaeology Coursebook* by Jim Grant, Sam Gorin, and Neil Fleming (2008; hereafter referred to as Grant).

Six frequency band word lists were created (Grant Frequency Band Lists; GFBL). The first contained words occurring 100 times or more in the text and the last contained only hapaxes (Table 1).

Table 1
Grant Frequency Band Lists (GFBL)

1st_100+.txt
2nd_50-99.txt
3rd_20-49.txt
4th_10-19.txt
5th_2-9.txt
6th_hapaxes.txt

Analysis

Readability

The readability scores across all three measures were consistent (Table 2). They showed the Orna text (FKRI = 54.3, FKGL = 9.8, SMOG = 12.1, Rank = 1) to be the easiest to read of the three texts and the McIntosh text (FKRI = 42.2, FKGL = 12.2, SMOG = 13.9, Rank = 3) to be the most difficult.

Lexical Diversity

Comparisons of lexical diversity indices (Table 3) showed the Orna text to have the lowest scores (TTRCI = 0.91, TTAI = 0.91, Rank = 1) and the McIntosh text to have the highest (TTRCI = 1.3, TTAI = 1.3, Rank = 3). Thus, the Orna text has the highest degree of vocabulary recycling out of the three texts. This can

Table 2
Readability Rankings

	Flesch-Kincaid Readability Index	Rank*	Flesch- Kincaid Grade Level	Rank*	SMOG	Rank*
Orna	54.2543	1	9.81163	1	12.1441	1
Barnes	50.8511	2	10.5046	2	12.7272	2
McIntosh	42.1909	3	12.1604	3	13.9154	3

* 1 = easiest

Table 3
Lexical Diversity Rankings

	TTRC Index	Rank*	TTA Index	Rank*
Orna	0.911	1	0.908	1
Barnes	1.123	2	1.083	2
McIntosh	1.298	3	1.265	3

* 1 = least diverse

be further confirmed with a bird’s-eye view of the type-token curves of the three texts (Figure 1). It is worth noting that the McIntosh text is written in a format that includes many detailed captions on cultural sub-topics. This variety in topic is likely what is inflating the lexical diversity of this text.

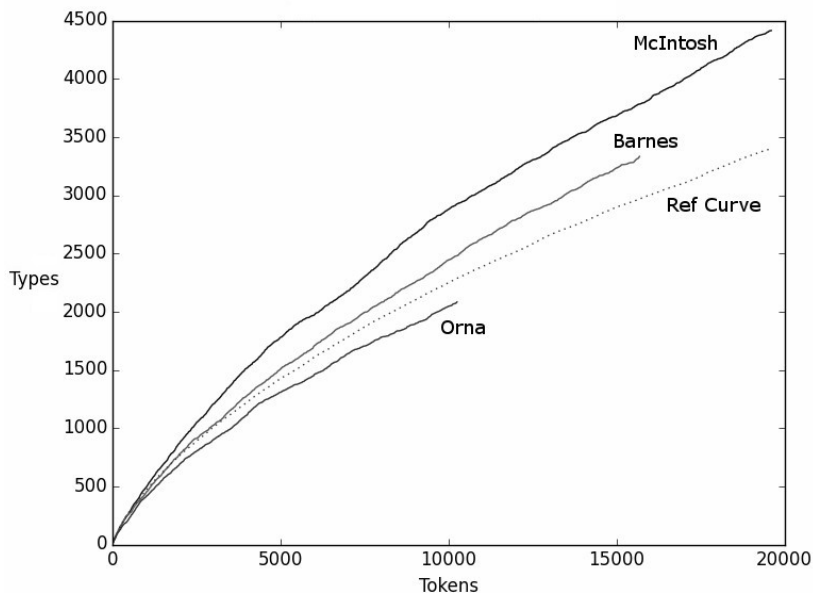


Figure 1. Type-token curve comparison. The text with the steepest curve has the highest lexical diversity.

Vocabulary Coverage

The Orna text showed the widest coverage of words from the GSL (Token% = 84.1, Rank = 1) and GFBL (Token% = 78.9, Rank = 1), while the McIntosh text had lower coverage (Token% = 77.1, Rank = 3; GFBL Token% = 67.9, Rank = 3; Table 4). However, with regard to the AWL, it was the McIntosh text that had the widest coverage (Token% = 5.9, Rank = 1, as compared with the Orna text's Token% = 4.3, Rank = 2, and the Barnes text's Token% = 3.9, Rank = 3). This is in line with the slightly more academic tone that, on closer reading, is noticeable in the McIntosh text.

Overall Ranking

The ranks for each textbook under each text analysis score were tallied for a total “best-candidate” overall ranking (Table 5). As can be seen, the Orna text (Overall Ranking = 8, where the lowest score is the best ranking) bears out as the most appropriate text for the students given the five criteria mentioned above (natural English, accessibility, vocabulary recycling, GSL/AWL vocabulary, and content-related vocabulary). While the Orna text does not seem to have as wide a coverage of the AWL as the McIntosh text, this is understandable given its higher coverage of GSL and GFBL words.

It is worth noting that, while not officially part of this study, this ranking was

Table 4
Vocabulary Coverage

	GSL Coverage Token % (levels 1-2)	GSL Coverage Type % (levels 1-2)	Rank*	AWL Coverage Token %	AWL Coverage Type %	Rank*	GFBL Coverage Token % (levels 1-3)	GFBL Coverage Type % (levels 1-3)	Rank*
Orna	84.13	63.26	1	4.26	8.99	2	78.85	35.92	1
Barnes	80.67	55.76	2	3.93	8.29	3	72.28	25.78	2
McIntosh	77.06	51.00	3	5.91	11.00	1	67.85	20.86	3

* 1 = easiest

Table 5
Overall Ranking of Texts

	Readability		Lexical Diversity		Vocabulary Coverage			Overall Ranking*
	Flesch-Kincaid	SMOG	TTRC Index	TTA Index	GSL	AWL	GFBL	
Orna	1	1	1	1	1	2	1	8
Barnes	2	2	2	2	2	3	2	15
McIntosh	3	3	3	3	3	1	3	19

* Lowest score is best candidate

later confirmed by students. When eventually given samples of these texts, they generally tended to agree that the McIntosh text was considerably more difficult to read than the Orna text.

Conclusion

In this study, text analysis measures were used to aid in the selection of a content textbook for an intermediate level themed university EFL class. Scores for the three texts showed a great deal of rank consistency across the different text analysis measures and could easily be tallied into an overall ranking that indicated the most appropriate candidate textbook for the class in question. Furthermore, the fact that there was such a high degree of rank consistency across the different measures lends a good deal of confidence to the selection of the textbook.

Notes

1. Subsequent deeper examination of, and experimentation with, this particular measure indicates possible flaws in its construction. However, the results for the TTRCI and visual agreement seen in the raw type-token curves are enough to confirm the textbook ranking from the point of view of lexical diversity.

References

- Anthony, L. (2014). AntWordProfiler (Version 1.4.0) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from <http://www.laurenceanthony.net/>
- Barnes, T. (2004). *Archaeology (Kingfisher Knowledge)*. Boston, MA: Kingfisher Publications.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Grant, J., Gorin, S., & Fleming, N. (2008). *The archaeology coursebook—An introduction to themes, sites, methods, and skills*. New York, NY: Routledge.
- Kincaid, J., Fishburne, R., Rogers, R., & Chissom, B. (1975). Derivation of new readability formulas (automated readability index, fog count, and Flesch reading ease formula) for Navy enlisted personnel. Research Branch Report 8-75. Memphis, TN: Naval Air Station Memphis.
- McIntosh, J. (2000). *Archaeology (Eyewitness Books)*. New York, NY: Dorling Kindersley Publishing, Inc.
- McLaughlin, G. (1969). SMOG grading – A new readability formula. *Journal of Reading*, 12(8), 639-646.
- Orna-Orntstein, J. (2002). *Archaeology – Discovering the past*. New York, NY: Oxford University Press.
- Silverstein, J. (2009). Readability (Version 1.0) [Computer Software]. Retrieved from <http://sourceforge.net/projects/readability2/>
- Thomas, D. (2016). Using type-token area indices to assess lexical diversity. In B. Kramer, I. Custance, C. Denison, & S. Porritt (Eds.), *Vocab@Tokyo 2016 Conference handbook: Current trends in vocabulary studies* (pp. 205-206). Retrieved from http://vli-journal.org/vocabattokyo/vocabattokyo_handbook_2016.pdf
- West, M. (1953). *A general service list of English words*. London, England: Longman, Green and Co.

Author bio

Dax Thomas has worked as an EFL instructor in Japan for nearly twenty years. His interests lay in corpus linguistics, vocabulary, and CLIL. He is a graduate of the University of Birmingham MA TESL/TEFL program and is currently an assistant professor in the Centre for Liberal Arts at Meiji Gakuin University. thomas@gen.meijigakuin.ac.jp

Received: November 23, 2017

Accepted: November 10, 2018