

Lexical Coverage and Readability in an ESL Reading Textbook: A Case Study of Reading Explorer 2 (Third Edition)

Bich Nguyen

Abstract

Vocabulary knowledge has long been recognized as playing a critical role in reading comprehension. Lexical coverage—the percentage of words in a text that a reader already knows—is a key predictor of text difficulty and, by extension, of reading comprehension. Understanding lexical coverage and readability can help teachers and materials developers select texts appropriate for their students. This study examines the lexical coverage and readability of an ESL reading textbook, Reading Explorer 2 (Third Edition), with the aim of offering practical guidance for textbook and material selection in reading classrooms.

Literature Review

Lexical Coverage in Reading Comprehension

Lexical coverage refers to the percentage of words in a text that readers or listeners already know. In reading, it is considered one of the primary factors affecting comprehension. Laufer (1989) examined the relationship between lexical coverage and reading comprehension scores by asking participants to mark the words they did and did not know in a set of reading passages, from which lexical coverage was calculated; participants then completed comprehension tests. The results showed that comprehension dropped significantly once coverage fell below 95%.

Hu and Nation (2000) conducted a related study examining how the density of unknown vocabulary in a text affects reading comprehension. Participants read different versions of a text and completed multiple comprehension measures, including a multiple-choice test and a cued written recall task in which they recorded what they remembered or understood. Comprehension scores were then compared across text versions. The results indicated that most learners required approximately 98% lexical coverage to comprehend a text without the aid of a dictionary, and that comprehension declined as the density of unknown words increased.

Building on this earlier work, Schmitt, Grabe, and Jiang (2011) investigated the relationship between the percentage of known vocabulary in a text and the level of comprehension achieved.

Participants completed a vocabulary test drawn from the reading passages before reading the texts and completing comprehension tests. The results showed no evidence of a discrete vocabulary “threshold” at which comprehension increased sharply; rather, the findings suggested that 98% coverage is a reasonable target for readers of academic texts.

Taken together, this body of research demonstrates a well-established relationship between lexical coverage and reading comprehension, underscoring the need for teachers and materials developers to consider the lexical coverage of textbooks and other reading materials. As Hu and Nation (2000) note, understanding lexical coverage allows teachers to select materials appropriate to different instructional purposes; the authors identify three types of reading in this regard:

- Intensive reading, in which learners work through texts containing unfamiliar language features but are supported by teacher guidance, discussion, dictionaries, glossaries, or in-text elaboration.
- Extensive reading for language growth, in which learners read texts containing some unfamiliar language features but can manage them with only minor interruptions to comprehension.
- Extensive reading for developing fluency, in which learners read texts with no unfamiliar language features so that reading is not interrupted by the need to process unknown language.

In sum, an understanding of lexical coverage can help teachers select the materials and textbooks best suited to their students and instructional goals.

Readability

Readability is another factor that influences reading comprehension. Traditionally, readability refers to the ease with which a text can be read and understood, a concept most clearly articulated by Dale and Chall (1949). They argued that readability involves more than surface-level features such as sentence length or word length; it also depends on factors such as vocabulary difficulty, sentence structure, content organization, and style. More recently, Zainurrahman, Yusuf, and Sukyadi (2024) found that text readability influences both students’ reading comprehension and reading time. Using the Friedman test, they reported that comprehension scores differed significantly across texts with different readability levels ($\chi^2 = 197.532$, $p < .001$), and that reading

time also varied significantly according to readability ($\chi^2 = 215.323$, $p < .001$), indicating that more difficult texts took longer to process and were associated with lower comprehension.

Carrell (1987), by contrast, cautioned against relying too heavily on readability formulas, arguing that such formulas capture only surface-level features of a text. In her study, she examined how readability and readers' background knowledge interact to influence comprehension. She found that background knowledge significantly improved comprehension overall; however, when a text was both linguistically difficult and unfamiliar in content, comprehension suffered considerably.

Taken together, these studies suggest that readability plays an important role in predicting reading comprehension, and that text difficulty should be considered alongside lexical coverage when selecting reading textbooks and materials.

Method

Research Design

This study employs a quantitative descriptive research design to examine the lexical coverage and readability of an ESL reading textbook. The analysis focuses on Reading Explorer 2 (Third Edition), published by National Geographic Learning. Each of the textbook's twelve chapters was analyzed individually, and the results were compared across chapters to identify trends in readability and lexical coverage over the course of the book.

Materials

The material analyzed in this study is Reading Explorer 2 (Third Edition), a textbook widely used in ESL/EFL reading courses. The textbook comprises twelve chapters, each built around one main reading passage; these twelve passages served as the units of analysis because they represent the core reading input provided to learners throughout the book. All passages were converted to plain text, with images, headings, and proper names removed to ensure compatibility with the analysis tools.

Procedure

Lexical Coverage Analysis

Lexical coverage was analyzed using AntWordProfiler, a corpus analysis tool that calculates the percentage of words in a text belonging to different frequency bands. The tool classifies words into the first 1,000 word families (K1), the second 1,000 word families (K2), Academic Word List (AWL) items, and off-list words. The analysis proceeded in three steps: (1) compiling the corpus by scanning the textbook chapters, (2) cleaning the text by removing images, diagrams, and other non-textual elements, and (3) generating word-frequency lists with AntWordProfiler for comparison across chapters.

Hypothesis. The classification of textbook readability into independent, instructional, and frustration levels based on lexical coverage draws on two research traditions rather than a single study. The three reading levels originate in L1 reading theory (e.g., Betts, 1946), which distinguishes between texts that readers can handle independently, with instruction, or only with difficulty. Subsequent L2 vocabulary research (e.g., Laufer, 1989; Hu & Nation, 2000; Nation, 2001) empirically demonstrated that readers typically require approximately 95% lexical coverage for independent comprehension, with comprehension declining sharply below this threshold. Researchers later aligned these lexical thresholds with the established reading levels, producing the framework used in this study: $\geq 95\%$ coverage corresponds to high (independent) readability, 75–95% to medium (instructional) readability, and $< 75\%$ to low (frustration) readability.

Readability Analysis

Text readability was assessed using an online readability analyzer that calculates several established readability indices, including the Flesch Reading Ease formula, which estimates text difficulty from surface-level linguistic features such as sentence length, word length, syllable count, and the proportion of unfamiliar words. Readability scores were calculated separately for each of the twelve chapters to capture variation in text difficulty across the textbook.

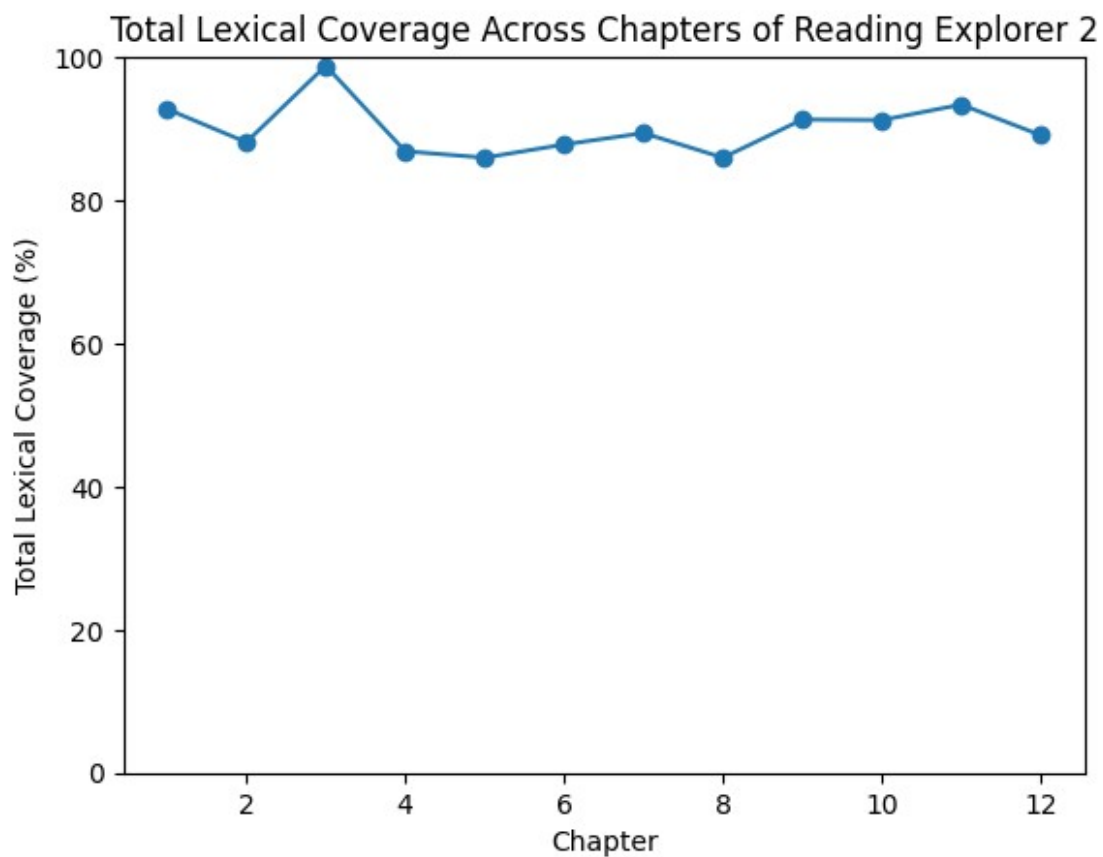
Data Analysis Procedures

Lexical coverage and readability scores obtained from the tools described above are presented in tables and graphs. Descriptive statistics were used to summarize patterns across chapters, and lexical coverage results were compared with readability scores. The analysis aimed to determine (a) whether the textbook chapters meet established lexical coverage thresholds for adequate

comprehension, (b) how readability and lexical coverage vary across chapters, and (c) whether readability scores are consistent with lexical coverage levels.

Results and Discussion

Lexical Coverage



Graph 1. Lexical coverage across the twelve chapters.

As shown in Graph 1 and Table 2, lexical coverage across the twelve chapters ranged from 85.96% to 98.81%, with most chapters falling between approximately 86% and 93%. According to the framework outlined above, most chapters therefore fall below the 95%–98% thresholds identified in previous research for independent reading (Laufer, 1989; Hu & Nation, 2000), placing them at the instructional level; that is, students are likely to need teacher assistance and guidance to fully comprehend the textbook.

The finding that most chapters do not reach the 95% threshold suggests that learners using this textbook may encounter lexical barriers to comprehension, particularly if the text is used for

independent or extensive reading. These chapters may instead be better suited to intensive reading, in which teacher support, vocabulary pre-teaching, or glosses are provided—consistent with Hu and Nation’s (2000) classification of reading types, in which texts with lower lexical coverage are better suited to guided instructional contexts.

Lexical coverage was also relatively consistent across chapters, suggesting that the textbook maintains a stable balance of known and unknown vocabulary throughout. This consistency may help students avoid the frustration associated with sharp fluctuations in lexical difficulty between chapters.

Chapter	Total Coverage	K1	K2	AWL	Off-list
1	92.76%	79.17%	9.06%	4.53%	7.25%
2	88.12%	75.00%	7.13%	5.85%	11.87%
3	98.81%	78.49%	7.29%	3.93%	10.29%
4	86.86%	74.08%	6.76%	6.04%	13.13%
5	85.96%	74.58%	7.25%	4.13%	14.04%
6	87.80%	75.86%	7.69%	4.25%	12.20%
7	89.40%	78.06%	6.74%	4.60%	10.60%
8	85.96%	75.66%	6.49%	3.81%	14.04%
9	91.29%	77.41%	8.38%	5.50%	8.71%
10	91.21%	77.14%	8.19%	5.88%	8.79%
11	93.38%	81.36%	6.43%	5.59%	6.61%
12	89.10%	76.82%	8.38%	3.90%	10.90%

Table 2. Lexical coverage data from AntWordProfiler.

As shown in Table 2, K1 words account for the largest proportion of vocabulary in every chapter, ranging from 75% to 81%, indicating that the textbook relies heavily on high-frequency vocabulary. AWL words, by contrast, typically range from 3% to 6%, suggesting that learners are regularly, though modestly, exposed to academic vocabulary.

Readability

Chapter	Flesch Reading Ease
---------	---------------------

1	71.03
2	70.25
3	76.27
4	68.29
5	69.10
6	72.70
7	66.55
8	68.42
9	68.60
10	67.41
11	69.47
12	77.30

Table 3. Flesch Reading Ease scores by chapter.

The Flesch Reading Ease scores for the twelve chapters ranged from 66.55 to 77.30, placing the text overall in the “standard” to “fairly easy” reading level. Scores above 60 generally indicate that a text is accessible to readers with at least a seventh- to eighth-grade reading level, suggesting that the textbook is accessible to a broad audience.

Most chapters clustered in the high 60s to low 70s, reflecting a fairly consistent readability level throughout the book. This consistency is beneficial because it allows readers to maintain comprehension without frequently adjusting to shifts in text difficulty between chapters. The modest variation that does exist suggests that some sections may be slightly more demanding, possibly owing to their content or the inclusion of more technical explanations.

Correlation Between Lexical Coverage and Readability

Overall, the chapters exhibited relatively high lexical coverage, ranging from 86.86% to 98.81%, suggesting that the textbook is generally at the instructional level. This finding indicates that students are likely to understand most of the text but will still benefit from teacher instruction.

In terms of readability, the Flesch Reading Ease scores ranged from 66.55 to 77.30, indicating that the book is fairly easy to understand. Texts in this range are generally accessible to teenage and adult readers, allowing for smooth reading without excessive cognitive effort.

Taken together, the lexical coverage and readability findings for this textbook do not conflict: both measures indicate that the textbook is appropriate for classroom use, provided it is supported by teacher instruction.

Conclusion

This study examined the lexical coverage and readability of Reading Explorer 2 (Third Edition) and found that most chapters fall at the instructional level. Lexical coverage ranged from 86.86% to 98.81%, indicating that students can understand much of the text but still require teacher support. The Flesch Reading Ease scores, ranging from 66.55 to 77.30, indicated that the text is fairly easy to read, although guidance remains beneficial.

These findings suggest that both lexical coverage and readability are important considerations when selecting reading materials for ESL/EFL learners. Teachers should attend to these factors to ensure that textbooks match their students' proficiency levels and learning goals. By considering lexical coverage and readability together, teachers can select materials that support comprehension, vocabulary growth, and reading-skill development while minimizing learner frustration. This study underscores that the careful selection of textbooks and materials, combined with appropriate teaching support, is central to helping students succeed.

References

- Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, 21(3), 461–481. <https://doi.org/10.2307/3586352>
- Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, 26(1), 19–26.
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Macmillan.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Schmitt, N., Grabe, W., & Jiang, X. (2011). The percentage of words known in a text and reading comprehension. *Language Learning*, 61(1), 75–95. <https://doi.org/10.1111/j.1467-9922.2010.00613.x>

Zainurrahman, Z., Yusuf, F. N., & Sukyadi, D. (2024). Readability and students' reading comprehension. *Journal of Language Teaching and Research*, 15(2), 210–222. <https://doi.org/10.17507/jltr.1502.05>