QUBAHAN EDUCATIONAL JOURNAL
VOL. 2, NO. 4, Sep 2025
https://doi.org/10.48161/qej.v2n4a10

The Relationship between Vocabulary Knowledge and Reading Comprehension among EFL Students

Nolwazi Cynthia Dlamini[1*], Jorge Isaac Torres Manrique[2]

[1] Gender and sexuality Studies, College of Human Sciences, University of South Africa, Pretoria, South Africa;
[2] Praeeminentia Iustitia Interdisciplinary School of Fundamental Rights, Catolica Santa Maria University, Peru;

* **Corresponding author:** lwazinduama1@gmail.com.

ABSTRACT: This quantitative correlational and predictive study examines the relationship between vocabulary knowledge and reading comprehension among university English as a Foreign Language (EFL) students. The study addresses an international academic-reading problem: university learners increasingly must process disciplinary, digital, and assessment-based texts in English, while vocabulary instruction in many EFL settings remains list-based, exam-oriented, and weakly connected to academic reading practices. We collected data from 318 undergraduate EFL learners using a vocabulary knowledge battery, a timed word-recognition task, a contextual guessing task, an academic vocabulary subtest, a vocabulary-depth subtest, and a reading comprehension test covering literal, inferential, vocabulary-in-context, and global comprehension across academic and non-academic passages. Reliability analysis indicated acceptable internal consistency for the main measures. Pearson correlation showed a strong positive relationship between composite vocabulary knowledge and reading comprehension ($r = .68$, $p < .001$). Multiple regression showed that vocabulary size, academic vocabulary, contextual guessing, word recognition, and vocabulary depth jointly explained 58% of the variance in reading comprehension ($R^2 = .58$; adjusted $R^2 = .57$). Vocabulary size was the strongest individual predictor, followed by academic vocabulary and contextual guessing. The study contributes a multidimensional lexical-reading prediction model showing that EFL reading difficulty is closely associated with lexical breadth, lexical quality, processing speed, academic vocabulary, and strategic contextual inference. Pedagogically, the findings support multidimensional vocabulary instruction as an evidence-based recommendation rather than a tested intervention effect.

Keywords: vocabulary knowledge; reading comprehension; EFL students; lexical quality; academic vocabulary.

I. INTRODUCTION

University EFL students increasingly face academic reading demands that go beyond typical classroom passages. They are expected to read textbooks, journal extracts, assessment passages, digital course materials, professional documents, and discipline-specific texts in English. In many contexts, however, vocabulary instruction remains list-based, translation-oriented, and disconnected from the academic reading tasks students must complete. The problem is not simply that learners know too few English words; their lexical knowledge may be too narrow, slow, shallow, or contextually inflexible to support academic reading.

The Lexical Quality Hypothesis explains the lexical basis of reading, arguing that skilled comprehension depends on precise word representations linking spelling, sound, grammatical behavior, and meaning [1]. When word representations are weak, readers spend more effort identifying words and have fewer cognitive resources for integration, inference, and monitoring comprehension. In EFL reading, this problem is intensified because learners may visually recognize a word but fail to access its meaning quickly or interpret it accurately in sentence and discourse context.

Lexical coverage research further shows that readers need a high proportion of known words in a text before higher-level comprehension strategies can operate effectively [2]. Vocabulary threshold arguments

similarly suggest that academic reading requires substantial lexical resources because frequent unknown words interrupt meaning construction and reduce inferential accuracy [3]. These perspectives are particularly relevant to university EFL learners, who may know high-frequency general words but remain weak in mid-frequency, low-frequency, and academic words that carry meaning in university-level texts.

Meta-analytic evidence confirms that vocabulary knowledge is strongly associated with L2 reading comprehension [4]. Related meta-analytic work among Chinese students also shows that vocabulary knowledge contributes meaningfully to reading comprehension, although effect sizes vary by age, assessment type, and comprehension measure [5]. These findings establish the importance of vocabulary, but they do not fully explain which lexical dimensions matter most when learners read both academic and non-academic texts.

A sharper model is therefore needed. Vocabulary size, vocabulary depth, academic vocabulary, word recognition, and contextual guessing represent different but related dimensions of lexical competence. Studies comparing breadth and depth show that vocabulary breadth may explain general comprehension performance, while depth becomes important for richer semantic processing [6]. Other work shows that recall and recognition are not equivalent indicators of lexical mastery, so vocabulary cannot be treated as a single, undifferentiated score [7].

The originality of the present study lies in its integrated model. Instead of confirming only that vocabulary and reading comprehension are related, the study examines vocabulary size, recognition speed, vocabulary depth, academic vocabulary, and contextual guessing together. This structure allows the analysis to identify which lexical dimensions most strongly predict reading comprehension and to translate the predictor pattern into carefully framed instructional recommendations rather than untested claims about intervention effects.

II. RESEARCH PROBLEM

This study addresses the lexical evidence gap in EFL reading comprehension. Poor reading performance is often described as a general comprehension weakness, but researchers do not always examine the specific lexical sources of difficulty systematically. A learner may fail to understand a text because of limited vocabulary size, slow word recognition, weak academic vocabulary, shallow semantic knowledge, or poor contextual guessing. These dimensions are related, yet they do not represent the same construct.

The first problem is measurement. Researchers often treat vocabulary knowledge as a single score, even though lexical competence includes breadth, depth, speed, academic word awareness, and strategic inference. A single vocabulary score can hide important differences among learners who know many words superficially, recognize words slowly, and infer unfamiliar meanings from context. A multidimensional measurement model is therefore necessary for explaining reading performance more precisely.

The second problem is pedagogical. In many EFL settings, vocabulary instruction continues to emphasize translation lists, isolated definitions, and short-term memorization for examinations. Such instruction may improve recognition of isolated words but may not develop academic word use, collocation knowledge, morphological awareness, or flexible contextual interpretation. Reading instruction may also overemphasize question-answer practice without strengthening the lexical resources needed to process the passage.

The third problem is an assessment problem. Reading failure may conceal lexical weaknesses. Students may appear unable to infer, summarize, or identify global meaning when the underlying difficulty is actually weak academic vocabulary, slow lexical access, or insufficient knowledge of word relations. This study addresses these problems by analyzing vocabulary knowledge as a multidimensional set of lexical predictors of EFL reading comprehension.

III. RESEARCH AIM, OBJECTIVES, AND QUESTIONS

1. RESEARCH AIM AND OBJECTIVES

The study aims to examine the relationship between vocabulary knowledge and reading comprehension among university EFL students and to identify the vocabulary dimensions that most strongly predict reading performance. To achieve this aim, the study combines multidimensional vocabulary measurement with academic and non-academic reading assessment through the following objectives:

- To measure EFL students' vocabulary knowledge through vocabulary size, word recognition, contextual guessing, vocabulary depth, and academic vocabulary indicators.
- To assess EFL students' reading comprehension performance across academic and non-academic texts, including literal, inferential, vocabulary-in-context, and global comprehension items.
- To determine the strength and direction of the relationship between vocabulary knowledge and reading comprehension achievement.
- To identify which vocabulary dimensions most strongly predict reading comprehension performance in a multiple regression model.
- To propose evidence-based recommended strategies for vocabulary-based reading instruction in EFL contexts.

2. RESEARCH QUESTIONS

- RQ1: What is the level of vocabulary knowledge among the university EFL students in the study?
- RQ2: What is the level of reading comprehension performance among the university EFL students in the study?
- RQ3: What is the relationship between vocabulary knowledge and reading comprehension performance among EFL students?
- RQ4: Which dimensions of vocabulary knowledge most strongly predict reading comprehension performance?
- RQ5: What recommended vocabulary-based pedagogical strategies may support reading comprehension development among EFL students?

IV. LITERATURE REVIEW

1. VOCABULARY KNOWLEDGE AS A MULTIDIMENSIONAL CONSTRUCT

Vocabulary knowledge is best understood as a multidimensional construct rather than as a simple count of known words. Schmitt's synthesis shows that vocabulary size and vocabulary depth are strongly connected but theoretically different [8]. Size refers to how many word families learners recognize, while depth refers to how well they know meanings, associations, collocations, derivations, polysemy, register, and constraints on use. This distinction matters because reading comprehension requires more than matching a written form with its most common meaning.

Receptive vocabulary size has been linked to multiple EFL language skills, but size alone does not explain reading success [9]. Learners may recognize a word in an isolated test item yet fail to use it while reading quickly, inferentially, or academically. Therefore, multidimensional assessment is necessary to distinguish breadth, quality, speed, and strategic use. The present study builds on this position by measuring vocabulary size, word recognition, contextual guessing, depth, and academic vocabulary separately.

2. VOCABULARY SIZE, LEXICAL COVERAGE, AND READING THRESHOLDS

Vocabulary size remains one of the strongest and most frequently studied predictors of reading comprehension. Masrai showed that high-, mid-, and low-frequency vocabulary bands contribute differently to L2 reading comprehension, which suggests that learners' reading difficulty may vary according to the lexical frequency profile of a text [10]. Size is often powerful because it increases text coverage and reduces the number of unknown words that interrupt meaning construction.

Academic vocabulary complicates the threshold argument because university texts contain abstract and information-dense words that do not always appear in general high-frequency lists. Lawrence and colleagues found meaningful relationships between academic vocabulary and reading comprehension across vocabulary measures [11]. Their later work showed that item features and reading proficiency jointly shape the academic

vocabulary-reading relationship, indicating that vocabulary difficulty cannot be reduced to frequency alone [12].

3. VOCABULARY DEPTH, ACADEMIC VOCABULARY, AND LEXICAL QUALITY

Vocabulary-depth research strengthens the argument that learners need qualitative lexical knowledge. Gnjidic, Kerestes, and Peti-Stanic demonstrated the value of measuring academic vocabulary breadth and depth through synonymy and collocational knowledge [13]. Ha also found relationships between receptive vocabulary dimensions and L2 academic listening and reading comprehension, showing that vocabulary knowledge interacts with academic language performance [14].

Vocabulary depth matters especially for inferential and academic comprehension because texts often require readers to interpret collocations, derivational forms, discourse markers, and context-specific meanings. Hu and colleagues showed that grammar and vocabulary may play direct and mediating roles in EFL reading comprehension, which means that vocabulary effects can interact with wider language knowledge [15]. Sidek and Rahim's cross-linguistic evidence likewise supports vocabulary as a key contributor to comprehension, but not as an isolated variable independent of reading processes [16].

4. EMPIRICAL EFL EVIDENCE AND CONTEXTUAL VARIATION

Empirical studies across EFL and ESL contexts generally confirm positive associations between vocabulary size and reading comprehension. Ibrahim, Sarudin, and Muhamad reported a significant relationship between vocabulary size and ESL reading comprehension [17]. Al Qunayeer found that vocabulary knowledge, reading comprehension, and proficiency level were related among Saudi EFL learners [18]. Manihuruk similarly reported a positive correlation between EFL students' vocabulary knowledge and reading comprehension in an Indonesian university context [19].

The strength of this relationship varies across contexts and measures. Beceren found that vocabulary size affected reading and listening comprehension among Turkish tertiary-level EFL learners [20]. Brooks, Clenton, and Fraser showed that vocabulary knowledge was highly important for reading comprehension among English as an additional language learners [21]. These findings support the robustness of the vocabulary-reading link, but they also suggest that context, task, and proficiency level influence how vocabulary knowledge is expressed in reading performance.

5. MORPHOLOGY, POLYSEMY, CONTEXTUAL GUESSING, AND PROCESSING SPEED

Vocabulary knowledge also includes morphological and inferential resources. Kieffer and Box demonstrated that derivational morphological awareness and academic vocabulary are connected to reading comprehension among linguistically diverse learners [22]. Zhang and Koda showed that morphological awareness and lexical inferencing ability contribute directly and indirectly to L2 vocabulary knowledge and reading comprehension [23]. These findings explain why researchers should include contextual guessing and vocabulary depth alongside vocabulary size.

Polysemous word knowledge is another depth-related dimension. Logan and Kieffer found that polysemous word knowledge contributes to reading comprehension among bilingual adolescents [24], and their later longitudinal work showed concurrent and longitudinal relationships between polysemous word knowledge and reading comprehension among Spanish-English bilingual students [25]. This evidence helps explain why a learner who knows a word's common meaning may still misinterpret it in academic or figurative contexts.

6. VOCABULARY DEVELOPMENT, ACADEMIC WORD LEARNING, AND EXTENSIVE READING

Repeated exposure, academic input, and intentional attention to word families strengthen vocabulary development. Coxhead and Boutorwick's longitudinal work in an English-medium international school context showed that learners' vocabulary develops in relation to the demands of EAL, mathematics, and science texts [26]. Cobb and Laufer's nuclear word family list further emphasizes the pedagogical value of high-frequency word-family knowledge [27].

Extensive reading provides another route for vocabulary growth because it offers repeated encounters with words in meaningful contexts. Nakanishi's meta-analysis found positive effects of extensive reading in second-language learning contexts [28]. Suk also reported positive effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition [29]. However, extensive reading alone may be insufficient for learners with weak academic vocabulary or limited inferencing skills, which supports the need for combined explicit and contextual vocabulary instruction.

7. CONTEXTUAL INFERENCE, ACADEMIC ASSESSMENT, AND REMAINING CONTRADICTIONS

Contextual guessing has pedagogical value, but it does not fully replace vocabulary knowledge. Hassanzadeh, Tamjid, and Ahangari found that instruction in lexical inference strategies was associated with incidental vocabulary acquisition and success in lexical inferencing [30]. Stevani and colleagues' work on contextual clues also supports the relevance of contextual processing in EFL reading comprehension [31]. These studies show that guessing from context may assist comprehension when learners possess enough surrounding lexical knowledge to make plausible inferences.

Academic vocabulary measurement remains a further challenge. Pecorari, Shaw, and Malmström developed a new academic vocabulary test, emphasizing the need for valid measures of academic vocabulary [32]. Masrai and Milton found that academic and general vocabulary knowledge both contribute to academic achievement [33]. Gu reported that vocabulary breadth and depth contributed to Chinese English learners' reading comprehension [34], while Baki found that vocabulary knowledge level affected EFL reading comprehension [35]. Manga Cetinavci showed that contextual factors shape word-meaning guessing [36]. Together, these studies reveal a consistent but nuanced pattern: vocabulary size is often the strongest predictor, but size alone is not sufficient for inferential, academic, or context-sensitive comprehension.

V. RESEARCH GAP AND NOVELTY

1. MULTIDIMENSIONAL VOCABULARY-MEASUREMENT GAP

Many vocabulary-reading studies rely on vocabulary size or a single receptive vocabulary score. This approach confirms that vocabulary matters but does not identify whether reading difficulty is more strongly associated with lexical breadth, recognition speed, academic vocabulary, depth of word knowledge, or contextual guessing. The present study addresses this gap by treating vocabulary knowledge as a set of interacting predictors rather than as one general score.

2. ACADEMIC VERSUS NON-ACADEMIC READING GAP

A second gap concerns the difference between academic and non-academic reading. Learners may perform adequately on general or literal passages while struggling with academic texts that require abstract vocabulary, dense noun phrases, logical connectors, nominalization, and inferential integration. The present study therefore examines reading comprehension across academic and non-academic passages and considers literal, inferential, vocabulary-in-context, and global comprehension subcomponents.

3. PEDAGOGY-SPECIFIC TRANSLATION GAP

A third gap concerns translating statistical findings into classroom recommendations. Prior studies often state that vocabulary instruction is important, but they rarely specify which vocabulary dimension should receive priority. This study does not simply confirm the vocabulary-reading relationship; it identifies which lexical dimensions matter most and translates the predictor pattern into cautious, evidence-based instructional recommendations based on predictive associations rather than demonstrated intervention effects.

VI. CONCEPTUAL AND THEORETICAL FRAMEWORK

The study draws on an integrated lexical-reading framework grounded in the Lexical Quality Hypothesis, reading comprehension theory, second language acquisition, and vocabulary acquisition. In this framework,

vocabulary size provides the breadth of known words, word recognition provides access speed, vocabulary depth provides semantic and relational quality, academic vocabulary provides access to specialized text meanings, and contextual guessing provides strategic processing of unfamiliar words.

The framework assumes these lexical dimensions relate to reading comprehension through lexical quality and processing efficiency. Stronger lexical representations may reduce cognitive load and support more accurate inferential and global comprehension. Because the design is correlational and predictive, the framework is interpreted as a predictive association model rather than a causal instructional model.

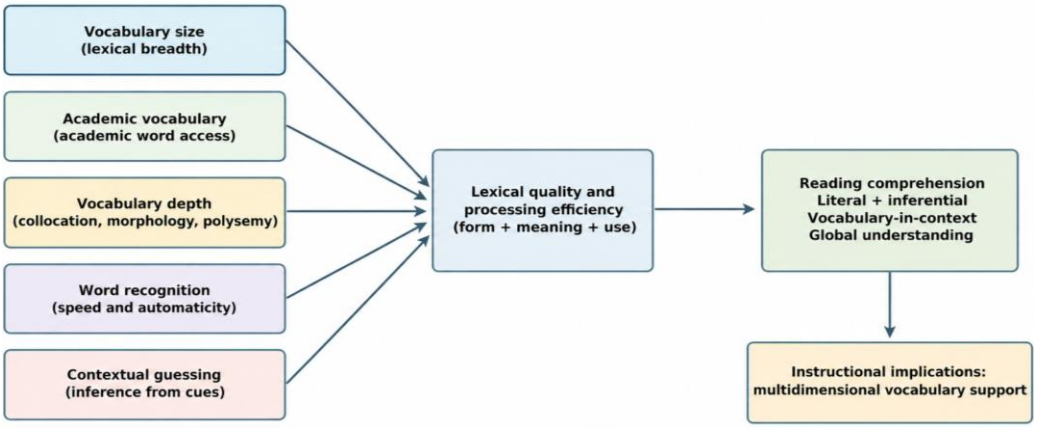


FIGURE 1. Conceptual framework linking vocabulary dimensions, lexical processing, reading comprehension, and instructional implications.

The model presents lexical dimensions as predictors of reading performance and links statistical interpretation to instructional implications.

Table 1. Theory-to-variable alignment for the vocabulary-reading framework.

Framework component	Variables/measures	Analytical role
Lexical Quality Hypothesis	Vocabulary depth, word recognition, lexical precision	Explains how higher-quality word representations are associated with more efficient comprehension.
Reading Comprehension Theory	Literal, inferential, global, and vocabulary-in-context comprehension	Explains reading as active meaning construction across lower- and higher-level processes.
Second Language Acquisition	Input, exposure, proficiency, lexical learning history	Explains how vocabulary grows through exposure, noticing, retrieval, and use.
Vocabulary Acquisition	Vocabulary size, academic vocabulary, contextual guessing	Explains how breadth, depth, and strategy combine in lexical development.
Pedagogical interpretation	Extensive reading, explicit teaching, word-family work	Links statistical associations to cautious instructional recommendations.

The alignment clarifies how each theoretical perspective informs the variables, measures, and analytical role used in the study.

VII. METHODOLOGY

1. RESEARCH DESIGN

The study used a quantitative correlational and predictive research design. This design was selected to determine the strength of the relationship between vocabulary knowledge and reading comprehension and to identify the vocabulary dimensions that predict reading outcomes. The design did not test an instructional intervention; instead, it examined naturally occurring differences in learners' lexical knowledge and reading performance.

The quantitative design was linked to pedagogical interpretation. Statistical results were used to identify which vocabulary dimensions require instructional attention, but the study does not claim that any recommended strategy has already produced measured instructional effects. This distinction protects the interpretation from unsupported causal claims.

2. PARTICIPANTS AND SAMPLING

The study involved 318 undergraduate EFL students enrolled in English-language courses at three university faculties. The study used stratified sampling to ensure representation across first-year, second-year, third-year, and fourth-year students and across lower-intermediate, intermediate, and upper-intermediate proficiency bands. Participants were aged between 18 and 24 years and had studied English formally for at least six years.

Inclusion criteria required participants to be enrolled in an EFL program, to have experience reading academic or general English texts, and to complete all vocabulary and reading tasks. We excluded participants who submitted incomplete test sheets, had missing demographic data, or completed the assessment in less than one-third of the expected time. The final sample included 318 valid cases after removing 14 incomplete responses from an initial pool of 332 participants.

Table 2. Participant profile, gender distribution, academic level, and proficiency criteria.

Group	n	Gender distribution	Academic level	Proficiency criteria
First-year EFL students	86	42 male; 44 female	Foundation/introductory courses	Placement score and course record: lower-intermediate to intermediate
Second-year EFL students	82	39 male; 43 female	Core language courses	Placement score and course record: intermediate
Third-year EFL students	78	38 male; 40 female	Academic reading/writing courses	Placement score and course record: intermediate to upper-intermediate
Fourth-year EFL students	72	37 male; 35 female	Advanced applied-language courses	Placement score and course record: upper-intermediate
Total	318	156 male; 162 female	University EFL learners	Mixed proficiency bands; all had at least six years of formal English study

The sample profile indicates balanced representation across academic levels and gender groups, with proficiency criteria based on placement scores and course records.

3. INSTRUMENTS

The instrument package included a vocabulary size test, a word-recognition task, a contextual guessing task, an academic vocabulary subtest, a vocabulary-depth subtest, a reading comprehension test, and a learner background questionnaire. The vocabulary size test measured receptive knowledge of high-frequency and mid-frequency word families. The word-recognition task measured speed and accuracy in identifying common and academic words. The contextual guessing task required learners to infer meanings of underlined words from sentence and paragraph context. The academic vocabulary subtest measured knowledge of words commonly used in academic texts, while the vocabulary-depth subtest measured synonyms, collocations, derivatives, and polysemous meanings.

The reading comprehension test contained four passages: two academic and two non-academic. Items measured main idea comprehension, supporting details, inference, vocabulary-in-context, reference, and global interpretation. We assessed instrument reliability using Cronbach's alpha and inter-rater agreement, and we pilot-tested with 38 students to check timing, item clarity, score distribution, and scoring consistency before main data collection.

Table 3. Data collection instruments, source/adaptation, scoring range, and reliability evidence.

Instrument	Construct measured	Source/adaptation	Scoring range	Reliability / scoring
Vocabulary size test	Receptive vocabulary breadth	Adapted from frequency-band vocabulary-size formats used in L2 vocabulary research	0-100 percentage score	alpha = .88
Word-recognition task	Speed and accuracy of lexical access	Timed recognition format adapted for high-frequency and academic words	0-100 accuracy-speed composite	alpha = .83
Contextual guessing task	Inference of unfamiliar meanings from context	Researcher-adapted short passages reviewed by two EFL specialists	0-100 rubric-based score	inter-rater agreement = .87
Academic vocabulary subtest	Knowledge of academic lexical items	Adapted from academic vocabulary assessment principles and course word lists	0-100 percentage score	alpha = .86
Vocabulary-depth subtest	Synonyms, collocations, derivations, polysemy	Adapted from vocabulary-depth and lexical-quality constructs	0-100 rubric/accuracy score	alpha = .84
Reading comprehension test	Literal, inferential, global, and vocabulary-in-context comprehension	Four researcher-adapted passages: two academic and two non-academic	0-100 total and subcomponent scores	alpha = .89

The instrument summary reports the construct measured, source or adaptation logic, scoring range, and reliability evidence for each measure.

4. DATA COLLECTION PROCEDURE

We collected data during regular university learning sessions with administrative permission and participant consent. Students first completed the background questionnaire, then the vocabulary-related measures, followed by the reading comprehension test. To reduce fatigue, we divided the assessment into two timed sessions separated by a short break. We coded test papers anonymously, and no score was included in course grades.

After data collection, responses were screened for missing values, careless response patterns, and timing anomalies. We entered scores into a spreadsheet and checked them twice against the original test sheets. Composite vocabulary knowledge was calculated from standardized scores for vocabulary size, word recognition, contextual guessing, academic vocabulary, and vocabulary depth. Reading comprehension was calculated as the total percentage score across the four reading passages and was also examined by subcomponent and text type.

5. STATISTICAL ANALYSIS

We analyzed the data using descriptive statistics, reliability analysis, Pearson correlation, and multiple regression. Descriptive statistics summarized the central tendency and dispersion of vocabulary and reading scores. Pearson correlation examined the strength and direction of relationships between each vocabulary dimension and reading comprehension. Multiple regression identified the unique contribution of vocabulary

size, word recognition, contextual guessing, academic vocabulary, and vocabulary depth to reading comprehension.

Regression diagnostics were reported to increase transparency. Multicollinearity was evaluated through tolerance and variance inflation factor (VIF) values for each predictor. VIF values ranged from 1.46 to 1.96, and tolerance values ranged from .51 to .68, indicating no problematic multicollinearity. Residual plots indicated acceptable linearity and homoscedasticity; the standardized residual distribution was approximately normal on the P-P plot; and standardized residuals remained within acceptable limits. We examined Cook's distance using the 4/n screening threshold (.013 for n = 318) and the conventional value of 1.0; the maximum Cook's distance was .011, so no influential case required removal. Outliers were retained because they did not distort model assumptions or exceed influence criteria. Confidence intervals were interpreted as the plausible range of each standardized predictor effect in the regression model.

6. VALIDITY, RELIABILITY, AND ETHICS

Validity was strengthened through the use of established vocabulary and reading constructs, expert review of test items, pilot testing, balanced academic and non-academic passages, and transparent scoring procedures. Acceptable internal consistency values supported reliability for the vocabulary and reading measures. Two trained raters reviewed contextual-guessing and vocabulary-depth items requiring short constructed responses; they resolved disagreements through a scoring guide and discussion.

Ethical safeguards included informed consent, voluntary participation, anonymity, non-punitive assessment, secure storage of test sheets, and the right to withdraw before data anonymization. Individual scores were not shared with teachers or administrators, and results were reported only in aggregate form.

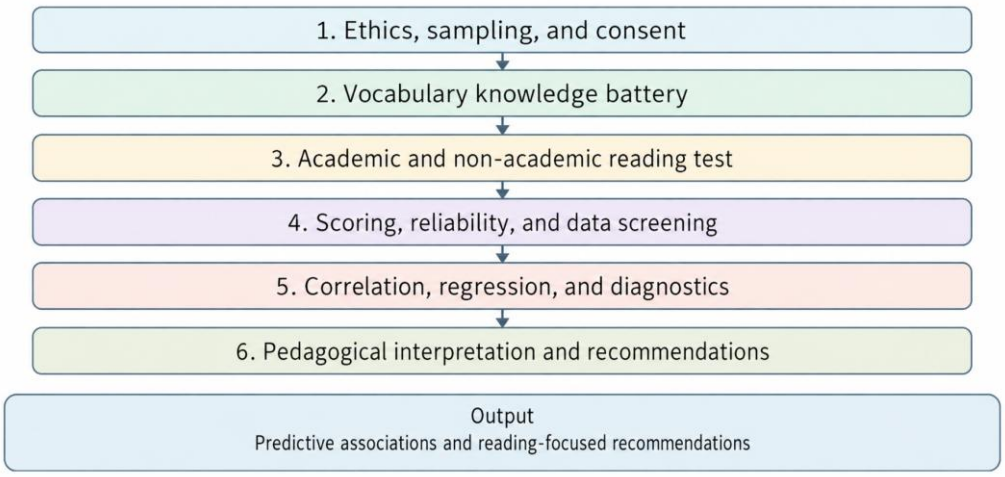


FIGURE 2. Quantitative correlational and predictive research workflow.

The workflow summarizes the sequence from ethics and sampling to assessment, scoring, statistical analysis, and pedagogical interpretation.

VIII. RESULTS AND FINDINGS

1. RQ1: VOCABULARY KNOWLEDGE LEVELS

Descriptive results showed that students performed strongest in word recognition and vocabulary size, while they performed weaker in academic vocabulary and contextual guessing. This pattern indicates that many learners possessed basic lexical recognition ability but had difficulty using vocabulary flexibly in academic and inferential reading contexts.

2. RQ2: READING COMPREHENSION PERFORMANCE

Reading comprehension performance was moderate overall. Students performed best on literal comprehension and non-academic passages, but their scores were lower on inferential comprehension, vocabulary-in-context items, and academic passages. This difference supports separating general reading performance from the lexical demands of academic reading.

Table 4. Descriptive statistics for vocabulary dimensions, reading subcomponents, and text-type performance.

Measure	Mean	SD	Minimum	Maximum	Interpretation
Vocabulary size	68.4	12.7	34	94	Moderate breadth with wide individual variation
Word recognition	71.8	11.4	38	96	Generally adequate recognition but uneven speed
Contextual guessing	62.5	13.6	28	91	Weakest lexical-strategy dimension
Academic vocabulary	59.7	14.2	25	90	Academic words remained a major difficulty
Vocabulary depth	63.1	12.9	30	92	Moderate knowledge of relations and word use
Reading comprehension total	65.9	13.8	31	93	Moderate overall reading achievement
Literal comprehension	72.4	11.6	39	96	Strongest reading subcomponent
Inferential comprehension	61.8	14.5	24	90	Weaker higher-level comprehension
Vocabulary-in-context	60.9	14.1	26	89	Sensitive to contextual lexical processing
Global comprehension	64.7	13.3	29	92	Moderate ability to integrate text meaning
Academic passage comprehension	61.5	14.6	27	90	Lower than non-academic passage performance
Non-academic passage comprehension	70.3	12.4	35	95	Relatively stronger text-type performance

The descriptive profile distinguishes overall vocabulary and reading performance from specific reading subcomponents and text-type outcomes.

3. RQ3: VOCABULARY-READING RELATIONSHIPS

Correlation analysis showed positive and statistically significant relationships between all vocabulary dimensions and reading comprehension. Composite vocabulary knowledge was strongly related to reading comprehension ($r = .68$, $p < .001$). Vocabulary size had the strongest bivariate association with reading comprehension, followed by academic vocabulary, contextual guessing, vocabulary depth, and word recognition. The correlation pattern also shows that vocabulary dimensions were mutually related but not identical, supporting the multidimensional conceptual framework.

Table 5. Pearson correlation matrix among vocabulary dimensions and reading comprehension; **p < .001.

Variable	1	2	3	4	5	6
1. Vocabulary size	1.00					
2. Word recognition	.54**	1.00				
3. Contextual guessing	.49**	.46**	1.00			
4. Academic vocabulary	.62**	.50**	.55**	1.00		
5. Vocabulary depth	.57**	.44**	.58**	.60**	1.00	
6. Reading comprehension	.64**	.52**	.58**	.61**	.55**	1.00

The matrix confirms strong positive associations while showing that the vocabulary dimensions are related but not interchangeable.

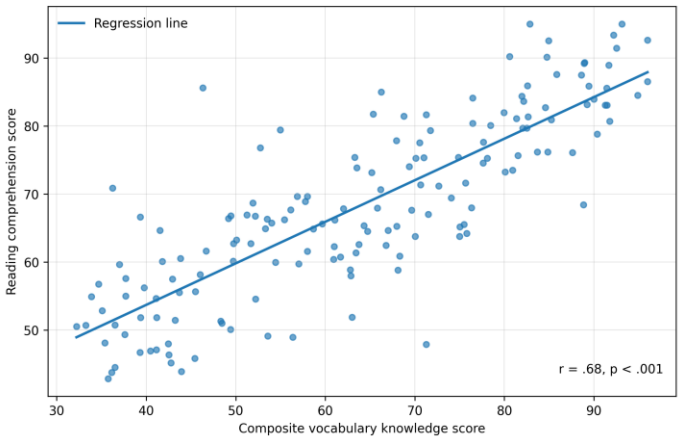


FIGURE 3. Predicted relationship between composite vocabulary knowledge and reading comprehension.

The scatterplot shows a clear positive trend between composite vocabulary knowledge and reading comprehension scores.

4. RQ4: PREDICTIVE MODEL AND DIAGNOSTIC RESULTS

We conducted multiple regression to identify the unique predictive contribution of each vocabulary dimension. The model was statistically significant and explained 58% of the variance in reading comprehension ($R^2 = .58$; adjusted $R^2 = .57$). Vocabulary size was the strongest predictor, followed by academic vocabulary and contextual guessing. Word recognition and vocabulary depth also contributed significantly, although their unique associations were smaller after accounting for the other vocabulary dimensions.

Table 6. Multiple regression model predicting reading comprehension performance with tolerance and VIF diagnostics.

Predictor	Standardized beta	SE	95% CI	p	Tolerance	VIF	Interpretive conclusion
Vocabulary size	.36	.05	[.27, .45]	< .001	.56	1.78	Strongest predictor; indicates the importance of lexical breadth.
Academic vocabulary	.29	.05	[.20, .38]	< .001	.52	1.92	Important for academic text comprehension.
Contextual guessing	.22	.06	[.11, .33]	< .001	.60	1.68	Associated with inference when unfamiliar words are encountered.
Word recognition	.18	.05	[.09, .27]	.002	.68	1.46	Faster access is associated with lower processing load.
Vocabulary depth	.16	.06	[.04, .28]	.011	.51	1.96	Richer word knowledge is associated with deeper comprehension.
Model summary	R^2 = .58; adjusted R^2 = .57						Diagnostics acceptable: VIF < 2.0; Cook's distance max = .011; no influential case removed.

The diagnostic values indicate acceptable multicollinearity and influence conditions, supporting the stability of the regression estimates.

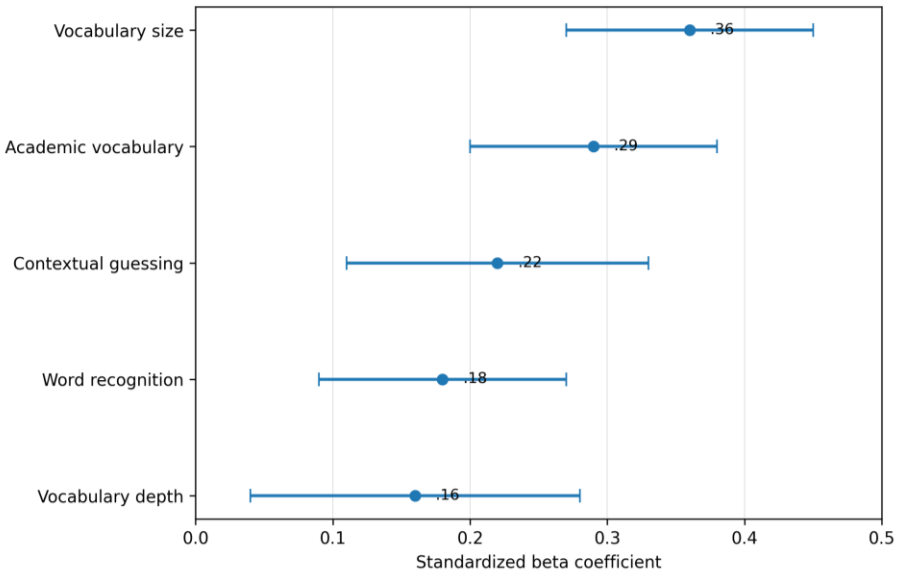


FIGURE 4. Standardized regression coefficients for vocabulary predictors of reading comprehension. The coefficient plot visualizes the standardized beta values and confidence intervals reported in Table 6.

5. RQ5: PEDAGOGICAL RECOMMENDATIONS

The findings support vocabulary-focused instructional recommendations, but these recommendations should not be interpreted as tested intervention effects. They identify areas of instructional priority derived from the observed predictive associations.

Table 7. Pedagogical interpretation of vocabulary-related findings as instructional recommendations.

Pedagogical target	Evidence from findings	Recommended instructional response
Vocabulary size	Strongest predictor of reading comprehension	Build high-frequency and mid-frequency word families through planned recycling and graded reading.
Academic vocabulary	Strong predictor of academic passage performance	Teach academic word families, nominalizations, reporting verbs, logical connectors, and discipline-relevant terms.
Contextual guessing	Weakest mean score but significant predictor	Teach inference from context, morphology, syntax, and discourse cues using guided practice.
Word recognition	Significant role in reading efficiency	Use timed recognition practice, repeated reading, and lexical fluency tasks.
Vocabulary depth	Associated with inferential and global comprehension	Teach collocations, polysemy, synonyms, derivatives, register, and word use in context.

Classroom actions are presented as recommendations derived from predictive associations, not as confirmed intervention effects.

IX. DISCUSSION

The findings support the argument that vocabulary knowledge is a major lexical predictor of reading comprehension among university EFL students. Regarding RQ1, the descriptive results showed that learners' vocabulary knowledge was moderate but uneven across dimensions. Students performed better on word recognition and vocabulary size than on academic vocabulary and contextual guessing. This pattern is consistent with the claim that vocabulary size and depth should be treated as related but distinct dimensions rather than as one score [8].

Regarding RQ2, reading comprehension results showed moderate performance overall, with stronger literal comprehension and weaker inferential, vocabulary-in-context, and academic passage results. This pattern is theoretically meaningful because learners can often answer literal items by identifying explicit information, whereas inferential and academic comprehension require richer lexical representations, discourse integration, and academic word knowledge. The finding aligns with lexical quality theory, which connects precise word representations with efficient comprehension [1].

RQ3 addressed the relationship between vocabulary knowledge and reading comprehension. The strong correlation between composite vocabulary knowledge and reading comprehension is consistent with meta-analytic evidence in L2 reading research [4], [5]. It also supports lexical coverage and threshold explanations [2], [3]. However, the result should not be reduced to the statement that more words automatically produce better reading. The correlation shows a strong predictive association, while actual reading performance may also be influenced by grammar knowledge, general proficiency, background knowledge, reading strategy use, motivation, and reading speed.

RQ4 asked which vocabulary dimensions most strongly predict reading comprehension. Vocabulary size was the strongest predictor, but academic vocabulary and contextual guessing also had substantial predictive value. This combination explains an important contradiction in the literature: vocabulary size often predicts reading comprehension best because it increases text coverage, but it is not sufficient for inferential or academic comprehension. Academic reading requires knowledge of abstract words, logical connectors, nominalized forms, reporting verbs, and discipline-related vocabulary, while inferential comprehension requires contextual processing and the integration of meaning across clauses and paragraphs.

Although contextual guessing had the lowest descriptive mean, it remained a significant predictor in the regression model. This indicates that learners' ability to infer meanings from context may be especially important when they encounter unfamiliar words in academic passages. Nevertheless, contextual guessing should not replace vocabulary growth. It is more accurate to state that contextual inference may support reading when learners already know enough surrounding language to make plausible guesses.

The results also explain why reading instruction should not rely only on comprehension-question practice. If students repeatedly answer reading questions without strengthening the lexical resources needed to understand the passage, instruction may diagnose comprehension difficulty without addressing one of its main lexical sources. Vocabulary instruction should therefore be systematic, text-linked, and multidimensional. It should include explicit vocabulary teaching, academic word development, word-family study, collocation work, contextual inference practice, repeated reading, and extensive reading.

The study is cross-sectional, correlational, and predictive. It shows that vocabulary dimensions predict reading performance, but it does not prove that vocabulary instruction will automatically improve reading comprehension. Future longitudinal and intervention research is needed to test whether targeted instruction in vocabulary size, academic vocabulary, contextual guessing, word recognition, or vocabulary depth produces measurable reading gains

X. THEORETICAL CONTRIBUTIONS

The main theoretical contribution is a multidimensional lexical-reading prediction model for EFL reading comprehension. The model treats vocabulary size, academic vocabulary, contextual guessing, word recognition, and vocabulary depth as related but distinct lexical predictors rather than as one general vocabulary score.

The study contributes to lexical quality research by showing that depth and recognition speed remain relevant even when vocabulary size is included in the model. It therefore links lexical representation quality to the efficiency of reading comprehension.

The study contributes to academic vocabulary research by identifying academic vocabulary as the second strongest predictor after vocabulary size. This finding links academic word knowledge directly to university EFL reading demands.

The study contributes to strategy-oriented vocabulary research by showing that contextual guessing is associated with reading comprehension, especially where learners face unfamiliar words. This supports the inclusion of contextual inference as a lexical-processing dimension.

Finally, the study contributes to EFL reading theory by linking lexical breadth, lexical depth, processing speed, academic vocabulary, and contextual inference in a single predictive framework to guide future empirical testing.

XI. PRACTICAL IMPLICATIONS

For EFL teachers, the findings suggest integrating vocabulary instruction with reading lessons. Teachers should pre-teach essential vocabulary, recycle target words across texts, and design reading tasks that require learners to use words in context rather than memorize isolated definitions.

For students, the findings indicate that reading development may benefit from active vocabulary learning. Learners should combine extensive reading with vocabulary notebooks, word-family study, contextual guessing practice, collocation learning, and academic vocabulary review. These strategies are presented as pedagogical recommendations derived from the observed associations, not as demonstrated intervention effects.

For curriculum designers, reading courses should include systematic vocabulary progression. Course materials should identify high-frequency, mid-frequency, and academic vocabulary targets and ensure repeated encounters across academic and non-academic texts.

For assessment designers, reading tests should include vocabulary-in-context and academic vocabulary components. This would make assessment more sensitive to the lexical dimensions that are associated with comprehension performance.

XII. LIMITATIONS AND FUTURE RESEARCH

The study has several limitations. First, it was conducted in a university EFL context, limiting direct generalization to other institutions, age groups, or educational systems. Second, the study did not test an instructional intervention, so the findings should be interpreted as predictive associations rather than causal evidence. Third, the design was cross-sectional and did not track vocabulary growth or reading development over time.

Fourth, test familiarity may have influenced performance because some learners may have been more accustomed to multiple-choice reading tasks or vocabulary tests. Fifth, the regression model did not include grammar knowledge, reading speed, general proficiency, motivation, background knowledge, or reading strategy use as control variables. Sixth, the study did not include qualitative learner data, such as think-aloud protocols or interviews, that could show how students actually use vocabulary while reading.

Future research should use longitudinal designs to examine how vocabulary growth predicts later reading development. Intervention studies should test whether targeted instruction in academic vocabulary, contextual guessing, word recognition, or vocabulary depth improves reading comprehension. Mixed-methods studies using think-aloud protocols, eye-tracking, learner interviews, and classroom observation would provide stronger evidence of how students activate vocabulary knowledge during academic and non-academic reading.

XIII. CONCLUSION

This study examined the relationship between vocabulary knowledge and reading comprehension among university EFL students through a quantitative correlational and predictive design. The results showed a strong positive relationship between composite vocabulary knowledge and reading comprehension ($r = .68$, $p < .001$). The regression model explained 58% of the variance in reading comprehension, with vocabulary size, academic vocabulary, contextual guessing, word recognition, and vocabulary depth all making significant predictive contributions.

The central conclusion is that EFL reading comprehension depends on multiple lexical dimensions, not vocabulary size alone. Vocabulary size provides broad lexical coverage, academic vocabulary supports university-level text interpretation, contextual guessing assists processing of unfamiliar words, word recognition is associated with reading efficiency, and vocabulary depth supports richer semantic and

inferential comprehension. These results explain why vocabulary size is often the strongest predictor but insufficient for academic and inferential reading.

The study contributes a multidimensional lexical-reading prediction model that links vocabulary breadth, lexical quality, recognition speed, academic vocabulary, and contextual inference to EFL reading comprehension. Vocabulary instruction should therefore be multidimensional, explicit, contextualized, and connected to reading tasks. However, future intervention research is needed to test the instructional effectiveness of specific vocabulary-based reading strategies.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. C. Perfetti, "Reading ability: Lexical quality to comprehension," Scientific Studies of Reading, vol. 11, no. 4, pp. 357-383, 2007, doi: 10.1080/10888430701530730.

[2]. N. Schmitt, X. Jiang, and W. Grabe, "The percentage of words known in a text and reading comprehension," The Modern Language Journal, vol. 95, no. 1, pp. 26-43, 2011, doi: 10.1111/j.1540-4781.2011.01146.x.

[3]. I. S. P. Nation, "How large a vocabulary is needed for reading and listening?" The Canadian Modern Language Review, vol. 63, no. 1, pp. 59-82, 2006, doi: 10.3138/cmlr.63.1.59.

[4]. S. Zhang and X. Zhang, "The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis," Language Teaching Research, vol. 26, no. 4, pp. 696-725, 2022, doi: 10.1177/1362168820913998.

[5]. Y. Dong, Y. Tang, B. W.-Y. Chow, W. Wang, and W.-Y. Dong, "Contribution of vocabulary knowledge to reading comprehension among Chinese students: A meta-analysis," Frontiers in Psychology, vol. 11, Article 525369, 2020, doi: 10.3389/fpsyg.2020.525369.

[6]. M. Li and J. R. Kirby, "The effects of vocabulary breadth and depth on English reading," Applied Linguistics, vol. 36, no. 5, pp. 611-634, 2015, doi: 10.1093/applin/amu007.

[7]. B. Laufer and T. Aviad-Levitzky, "What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition?" The Modern Language Journal, vol. 101, no. 4, pp. 729-741, 2017, doi: 10.1111/modl.12431.

[8]. N. Schmitt, "Size and depth of vocabulary knowledge: What the research shows," Language Learning, vol. 64, no. 4, pp. 913-951, 2014, doi: 10.1111/lang.12077.

[9]. I. Miralpeix and C. Muñoz, "Receptive vocabulary size and its relationship to EFL language skills," International Review of Applied Linguistics in Language Teaching, vol. 56, no. 1, pp. 1-24, 2018, doi: 10.1515/iral-2017-0016.

[10]. A. Masrai, "Vocabulary and reading comprehension revisited: Evidence for high-, mid-, and low-frequency vocabulary knowledge," SAGE Open, vol. 9, no. 2, pp. 1-13, 2019, doi: 10.1177/2158244019845182.

[11]. J. F. Lawrence, A. M. Hagen, J. K. Hwang, G. Lin, and A. Lervåg, "Academic vocabulary and reading comprehension: Exploring the relationships across measures of vocabulary knowledge," Reading and Writing, vol. 32, no. 2, pp. 285-306, 2019, doi: 10.1007/s11145-018-9865-2.

[12]. J. F. Lawrence, R. E. Knoph, A. McIlraith, P. A. Kulesz, and D. J. Francis, "Reading comprehension and academic vocabulary: Exploring relations of item features and reading proficiency," Reading Research Quarterly, vol. 57, no. 2, pp. 669-690, 2022, doi: 10.1002/rrq.434.

[13]. V. Gnjidic, G. Kerestes, and A. Peti-Stanić, "Measuring Academic Vocabulary Breadth and Depth in Upper-Elementary School Students," Croatian Journal of Education, vol. 23, no. 4, pp. 1043-1085, 2021, doi: 10.15516/cje.v23i4.4145.
[14]. H. T. Ha, "Exploring the relationships between various dimensions of receptive vocabulary knowledge and L2 listening and reading comprehension," Language Testing in Asia, vol. 11, Article 20, 2021, doi: 10.1186/s40468-021-00131-8.
[15]. T.-C. Hu, Y.-T. Sung, H.-H. Liang, T.-J. Chang, and Y.-T. Chou, "Relative roles of grammar knowledge and vocabulary in the reading comprehension of EFL elementary-school learners: Direct, mediating, and form/meaning-distinct effects," Frontiers in Psychology, vol. 13, Article 827007, 2022, doi: 10.3389/fpsyg.2022.827007.
[16]. H. M. Sidek and H. A. Rahim, "The role of vocabulary knowledge in reading comprehension: A cross-linguistic study," Procedia - Social and Behavioral Sciences, vol. 197, pp. 50-56, 2015, doi: 10.1016/j.sbspro.2015.07.046.
[17]. E. H. E. Ibrahim, I. Sarudin, and A. J. Muhamad, "The relationship between vocabulary size and reading comprehension of ESL learners," English Language Teaching, vol. 9, no. 2, pp. 116-123, 2016, doi: 10.5539/elt.v9n2p116.
[18]. H. S. Al Qunayeer, "An investigation of the relationship between reading comprehension, vocabulary knowledge, and English language proficiency level of Saudi EFL learners," Advances in Language and Literary Studies, vol. 12, no. 2, pp. 59-65, 2021, doi: 10.7575/aiac.alls.v.12n.2.p.59.
[19]. D. H. Manihuruk, "The correlation between EFL students' vocabulary knowledge and reading comprehension: A case study at the English Education Department of Universitas Kristen Indonesia," Journal of English Teaching, vol. 6, no. 1, pp. 86-95, 2020, doi: 10.33541/jet.v6i1.1264.
[20]. S. Beceren, "Vocabulary size of Turkish EFL learners at the tertiary level and its impact on their reading and listening comprehension," Akdeniz Eğitim Araştırmaları Dergisi, vol. 14, no. 32, pp. 206-219, 2020, doi: 10.29329/mjer.2020.258.11.
[21]. G. Brooks, J. Clenton, and S. Fraser, "Exploring the importance of vocabulary for English as an additional language learners' reading comprehension," Studies in Second Language Learning and Teaching, vol. 11, no. 3, pp. 351-376, 2021, doi: 10.14746/ssllt.2021.11.3.3.
[22]. M. J. Kieffer and C. D. Box, "Derivational morphological awareness, academic vocabulary, and reading comprehension in linguistically diverse sixth graders," Learning and Individual Differences, vol. 24, pp. 168-175, 2013, doi: 10.1016/j.lindif.2012.12.017.
[23]. D. Zhang and K. Koda, "Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects," Reading and Writing, vol. 25, pp. 1195-1215, 2012, doi: 10.1007/s11145-011-9313-z.
[24]. J. K. Logan and M. J. Kieffer, "Evaluating the role of polysemous word knowledge in reading comprehension among bilingual adolescents," Reading and Writing, vol. 30, no. 8, pp. 1687-1704, 2017, doi: 10.1007/s11145-017-9745-1.
[25]. J. K. Logan and M. J. Kieffer, "Investigating the longitudinal and concurrent relationships between polysemous word knowledge and reading comprehension among Spanish-English bilingual middle school students," Reading and Writing, vol. 34, no. 2, pp. 301-335, 2021, doi: 10.1007/s11145-020-10071-z.
[26]. A. Coxhead and T. J. Boutorwick, "Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, maths, and science," TESOL Quarterly, vol. 52, no. 3, pp. 588-610, 2018, doi: 10.1002/tesq.450.
[27]. T. Cobb and B. Laufer, "The nuclear word family list: A list of the most frequent family members, including base and affixed words," Language Learning, vol. 71, no. 3, pp. 834-871, 2021, doi: 10.1111/lang.12452.
[28]. T. Nakanishi, "A meta-analysis of extensive reading research," TESOL Quarterly, vol. 49, no. 1, pp. 6-37, 2015, doi: 10.1002/tesq.157.
[29]. N. Suk, "The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition," Reading Research Quarterly, vol. 52, no. 1, pp. 73-89, 2017, doi: 10.1002/rrq.152.
[30]. Z. Hassanzadeh, A. Tamjid, and F. Ahangari, "The effect of lexical inference strategy instruction on Iranian EFL learners' incidental vocabulary acquisition and lexical inferencing success," Cogent Education, vol. 6, no. 1, Article 1614750, 2019, doi: 10.1080/2331186X.2019.1614750.
[31]. M. Stevani, M. S. Prayuda, D. W. Sari, S. M. Marianus, and K. E. Tarigan, "Evaluation of contextual clues: EFL proficiency in reading comprehension," English Review: Journal of English Education, vol. 10, no. 3, pp. 993-1002, 2022, doi: 10.25134/erjee.v10i3.7076.
[32]. D. Pecorari, P. Shaw, and H. Malmström, "Developing a new academic vocabulary test," Journal of English for Academic Purposes, vol. 39, pp. 59-71, 2019, doi: 10.1016/j.jeap.2019.02.004.
[33]. A. Masrai and J. Milton, "Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement," Journal of English for Academic Purposes, vol. 31, pp. 44-57, 2018, doi: 10.1016/j.jeap.2017.12.006.

[34]. T. Gu, "The effect of vocabulary knowledge on Chinese English learners' reading comprehension," International Journal of English Linguistics, vol. 7, no. 4, pp. 45-55, 2017, doi: 10.5539/ijel.v7n4p45.
[35]. R. B. Baki, "The impact of vocabulary knowledge level on EFL reading comprehension," International Journal of Applied Linguistics and English Literature, vol. 2, no. 2, pp. 85-89, 2013, doi: 10.7575/aiac.ijalel.v.2n.2p.85.
[36]. B. Manga Cetinavci, "Contextual factors in guessing word meaning from context in a foreign language," Procedia - Social and Behavioral Sciences, vol. 116, pp. 2670-2674, 2014, doi: 10.1016/j.sbspro.2014.01.633.

Appendices

1. *APPENDIX A. VOCABULARY TEST SAMPLE ITEMS*

The following sample items illustrate the item types used in the vocabulary knowledge battery; operational test items may be withheld to protect assessment security.

Appendix Table A1. Sample vocabulary-test item types.

Dimension	Sample item	Expected response/scoring
Vocabulary size	Choose the closest meaning of accumulate: (a) collect (b) divide gradually (c) avoid (d) forget	One point for correct recognition.
Word recognition	Timed decision: Is analyze an academic English word? yes/no	Accuracy is scored with response time in the composite.
Academic vocabulary	Complete the sentence: The study provides strong ____ for the argument. (evidence / weather / furniture / travel)	One point for academic word selection.
Contextual guessing	The findings were inconsistent; therefore, the researcher interpreted them cautiously. Meaning of inconsistent?	Credit for "not the same," "not stable," or equivalent contextual inference.
Vocabulary depth	Select the best collocation: make / conduct / do a study	Credit for conduct a study.

2. *APPENDIX B. READING-COMPREHENSION TEST BLUEPRINT*

Appendix Table B1. Reading-comprehension test blueprint.

Passage type	Number of passages	Items	Subcomponents	Score range
Academic	2	20	Literal, inferential, vocabulary-in-context, global comprehension	0-50
Non-academic	2	20	Literal, inferential, vocabulary-in-context, global comprehension	0-50
Total	4	40	Balanced coverage of lower- and higher-level comprehension	0-100

3. *APPENDIX C. CONTEXTUAL GUESSING SCORING GUIDE*

Appendix Table C1. Contextual guessing scoring guide.

Score	Criterion	Description
2	Accurate inference	Meaning is contextually appropriate and semantically precise.
1	Partially accurate inference	Meaning is generally related but incomplete, too broad, or weakly phrased.
0	Inaccurate or missing inference	Response contradicts the context, gives an unrelated meaning, or is blank.

4. *APPENDIX D. VOCABULARY-DEPTH SCORING RUBRIC*

Appendix Table D1. Vocabulary-depth scoring rubric.

Dimension	Full-credit evidence	Partial-credit evidence
Synonym knowledge	Correctly identifies the closest synonym in context.	Selects a semantically related but less precise option.
Collocation knowledge	Selects a natural and conventional word combination.	Selects a possible but less natural combination.
Derivational knowledge	Correctly transforms or recognizes word-family forms.	Recognizes the base word but not the accurate derived form.
Polysemy	Identifies the context-specific meaning of a multi-meaning word.	Gives a common meaning that is only partly compatible with the context.

5. APPENDIX E. RELIABILITY, VALIDITY, AND PILOT-TESTING DETAILS

Appendix Table E1. Reliability, validity, and pilot-testing procedures.

Quality-control stage	Procedure	Outcome/use
Expert review	Two EFL specialists reviewed item clarity, construct coverage, and academic appropriateness.	Items with ambiguous wording were revised before piloting.
Pilot testing	Thirty-eight students completed the instrument package before the main study.	We checked timing, item difficulty, and score spread.
Reliability analysis	Cronbach's alpha was calculated for objective scales; inter-rater agreement was calculated for constructed-response tasks.	All main measures reached acceptable reliability values.
Rater training	Two raters scored contextual guessing and vocabulary-depth responses using rubrics.	We resolved disagreements through discussion and clarification of the scoring guide.
Data screening	We checked missing values, response-time anomalies, standardized residuals, and Cook's distance.	No influential case exceeded removal criteria.