Journal of Linguistics & Cultural Studies

Article

Effectiveness of English Vocabulary and Discourse-analysis Skills in Reading Comprehension among Chinese Middle Schools

Huijun Niu[1,*]

[1] School of Foreign Languages, Zhaotong University, Zhaotong, China
* Correspondence: Huijun Niu, School of Foreign Languages, Zhaotong University, Zhaotong, China

Abstract: Vocabulary and discourse-analysis skills constitute the fundamental core components of reading comprehension. The present empirical study explored the effectiveness of these linguistic competencies in reading comprehension among Chinese middle school students. Conducted in Dalian City, Liaoning Province, China, the research assessed thirty middle school students using standardized tests for vocabulary, discourse-analysis skills, and reading comprehension. Correlation analyses revealed that vocabulary knowledge, inferential reasoning, and discourse-analysis skills are significantly relevant to reading comprehension outcomes. Furthermore, regression analysis indicated that these variables significantly improve and enhance students' overall reading comprehension proficiency. Path analysis demonstrated that the direct paths from vocabulary knowledge, inferential reasoning, text structure, and discourse-analysis skills to reading comprehension were highly significant. Additionally, discourse-analysis skills exhibited significant indirect effects on reading comprehension through mediating variables. The study concludes that reading comprehension is influenced by multiple elaborated factors: first, vocabulary knowledge and inferential reasoning serve as foundational skills that operate through their impact on discourse-analysis skills; second, text structure and discourse-analysis skills independently and jointly contribute to reading comprehension. Ultimately, reading comprehension instruction must prioritize vocabulary and discourse analysis skills to adequately prepare students for real-life encounters with complex texts across academic, professional, and personal settings. These findings provide valuable pedagogical implications for English language teaching methodologies in secondary education.

Keywords: vocabulary knowledge; discourse analysis; reading comprehension; inferential reasoning; secondary education

Received: 18 September 2026
Revised: 07 November 2026
Accepted: 19 November 2026
Published: 24 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Reading is the process of extracting operational information from a discourse, and this extraction process is dynamic rather than fixed. Effective information extraction requires readers to automatically encode textual linguistic symbols into mental phonological representations, access accumulated semantic knowledge stored in the mental lexicon, and integrate that knowledge with newly generated information from the ongoing reading process. Readers then synthesize prior and current information, internalizing it as part of their cognitive framework. A communicative approach to contemporary English language teaching emphasizes discourse comprehension, incorporating techniques such as prediction, information retrieval, holistic understanding, inference of stance and attitude, and contextual meaning derivation. Vocabulary plays a central role in reading comprehension, and research has consistently highlighted its significance across different learner populations. Studies conducted in English-as-a-foreign-language contexts have examined the interplay between vocabulary knowledge, discourse-level strategies, and reading outcomes among middle school learners. Other

investigations have focused on how discourse analysis abilities support reading comprehension, including word sense inference and incidental vocabulary acquisition. Recent work has explored instructional models that bridge lexical gaps to strengthen sustained comprehension and foster more consistent reading habits [1]. Core reading strategies—such as imaginative engagement, predictive reasoning, and logical inference—have been adapted for classroom implementation through carefully designed instructional language and lesson fragments aligned with contemporary pedagogical goals and curriculum materials. Such adaptations aim to ensure robust integration of theoretical principles with practical application. Further research has underscored the foundational role of vocabulary instruction in reading development, while also treating listening, speaking, reading, and writing as interconnected dimensions of comprehension.

However, a review of existing literature on English reading instruction for junior high school students reveals that many proposed pedagogical approaches prioritize the development of general reading ability and strategy use, while giving insufficient attention to the systematic enhancement of vocabulary knowledge and discourse analysis competencies. To address this gap, the present study investigates English reading comprehension among Chinese junior high school students through an applied lens, focusing specifically on the synergistic effects of vocabulary development and discourse analysis training. The aim is to support learners in strengthening their motivation for English reading, diversifying their reading approaches, and improving overall reading performance. Beyond linguistic competence, discourse analysis equips students to recognize cultural nuances embedded in English texts. Accurate interpretation of cultural context is essential for accessing deeper textual meaning, thereby offering a pathway to cultivate intercultural awareness alongside language learning. Integrating targeted vocabulary instruction within reading tasks supports lexical expansion and familiarity with varied expressions. Enhanced language proficiency enables more accurate text interpretation and more precise self-expression [2, 3].

2. Definition

2.1. Reading Comprehension

Reading comprehension is the ability to understand and interpret written text. It involves actively engaging with text to extract meaning and construct a coherent understanding of the information presented. This process encompasses multiple interrelated skills, including vocabulary knowledge, sentence structure analysis, logical reasoning, identification of main ideas, and drawing evidence-based conclusions. Effective reading comprehension integrates decoding—the capacity to recognize and interpret words—with higher-order comprehension—the capacity to grasp and synthesize the meaning of a text as a whole [4, 5]. Skilled readers achieve rapid and accurate word recognition, interpret lexical and contextual meaning efficiently, and continuously integrate new information with their evolving mental representation of the text. Research has explored various theoretical perspectives on reading comprehension, including schema-based models of processing, analyses of reading rate and its instructional implications, classroom-based strategies for enhancing motivation and achievement, construction-integration frameworks emphasizing the role of prior knowledge in discourse understanding, verbal protocol methodologies for investigating reader cognition, and evidence-informed practices for instructional development.

Reading comprehension is the ability to process written text, interpret its meaning, and integrate it with the reader's existing knowledge [6]. Specifically, comprehension constitutes a creative and multifaceted process grounded in four foundational language dimensions: phonology, syntax, semantics, and pragmatics.

2.2. Vocabulary

A rich vocabulary is fundamental for comprehending the meaning of words within a text. The greater the number of words a reader recognizes, the more effectively they can understand the text and establish meaningful connections among its elements [7].

Vocabulary encompasses both receptive and productive knowledge. Receptive vocabulary refers to learners' ability to understand a word upon hearing or seeing it, while productive vocabulary denotes their capacity to accurately use the word in speaking or writing. It is widely observed that learners typically acquire receptive knowledge first, with productive use emerging only after deliberate, conscious learning. Furthermore, receptive vocabulary gradually transitions into productive knowledge as learners develop heightened awareness of lexical items.

The relationship between receptive and productive vocabulary is illustrated in the following Figure 1:

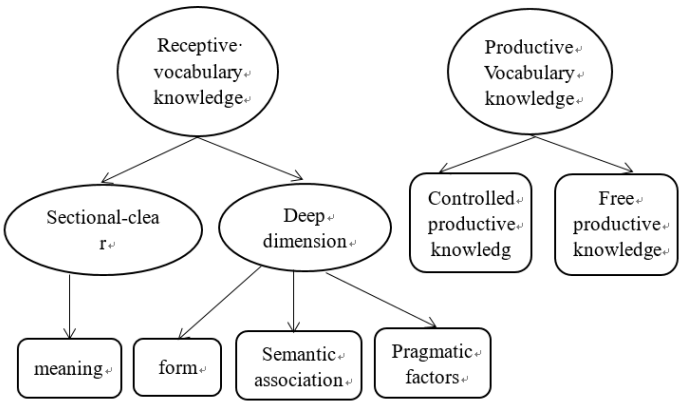

Figure 1. The Relations between Receptive and Productive Vocabulary (Zhong, 2012)

A critical stage in vocabulary development is the transition from receptive to productive knowledge, often regarded as the culmination of vocabulary acquisition. This progression begins with receptive knowledge, as most vocabulary is initially acquired through exposure rather than active production. Consequently, receptive vocabulary tends to exceed productive vocabulary in scope, offering insight into the relative size and developmental trajectory of productive knowledge. While receptive understanding emerges early, productive use generally requires intentional learning strategies. Some studies suggest that even modest gains in productive knowledge significantly reinforce receptive retention; however, other research indicates that this relationship may not be consistently directional [8].

A robust vocabulary is essential for interpreting written texts [9]. Insufficient lexical resources can impede comprehension of individual words and syntactic structures, thereby compromising overall textual understanding. Expanding vocabulary enables learners to perceive subtle semantic distinctions in English texts and enhances reading comprehension. To strengthen both lexical competence and discourse analysis skills, educators may employ diverse instructional approaches—including targeted vocabulary instruction, context-embedded reading tasks, and explicit teaching of discourse features. Consistent practice, sustained engagement with authentic English materials, and formative assessment further support the ongoing development of these competencies.

2.3. Discourse Analysis

Generally speaking, 'discourse' relates to the actual practice of speaking and writing [10]. Our use of the term is more precise: 'discourse' is defined as a set of interrelated texts and their associated practices of production, circulation, and reception that collectively constitute an object of knowledge. For example, a constellation of historical texts and institutional practices contributed to the emergence of certain psychological concepts in the nineteenth century. This perspective underscores that social reality is constituted through language-in-use, and that social interaction cannot be fully understood without attending to the discursive frameworks that confer meaning upon it. Therefore, the central

task of discourse analysis is to examine the dynamic relationship between discourse and the realities it helps shape.

Discourse analysis is the systematic study of how language is structured and organized beyond the sentence level. It focuses on how clauses and sentences are linked to form coherent, purposeful texts [11]. Insights into the functional organization of language date back to early twentieth-century linguistic inquiry, and the conceptual foundation for analyzing units larger than the sentence was established in mid-century scholarship. Intensive investigation into discourse structure gained momentum from the 1960s onward, with foundational work examining paragraph-level organization. Concurrently, research into narrative patterns in spoken accounts advanced theoretical models of cross-linguistic discourse parameters. Studies addressing textual cohesion—distinguishing unified texts from arbitrary sequences of sentences—emerged prominently within systemic functional frameworks during the same period. Further contributions explored how communicative functions influence grammatical choices at the clause level. In parallel, conversation analysts described recurrent structural features of spoken interaction, such as openings and closings, situating these within broader sociological inquiries into how people collaboratively construct social order through talk. Subsequent formal models of discourse structure, analogous to syntactic or morphological models, include hierarchical coherence frameworks and rhetorical structure theories that map relations among text segments.

Discourse-analysis skills involve understanding the structural architecture, logical progression, and cohesive mechanisms of written texts [12]. Developing these skills enables learners to identify central ideas, trace relationships among propositions, and follow the overall argumentative or expository flow. Such competence supports accurate interpretation of authorial intent, recognition of main ideas and supporting evidence, and awareness of rhetorical strategies—all contributing to enhanced reading comprehension. Discourse analysis entails examining how language operates across diverse contexts, with particular attention to textual organization, functional coherence, and contextual appropriateness. Cultivating these analytical capacities helps learners discern semantic connections, locate salient information, and navigate the logical architecture of a text, thereby improving their ability to interpret meaning and grasp the intended message.

Since the 1970s, discourse analysis has been increasingly integrated into foreign language pedagogy. It is widely recognized as a meaning-making process in which readers draw upon prior knowledge and experience to reconstruct the conceptual and affective dimensions embedded in a text [10, 13]. This reconstructive engagement invites readers to move beyond personal assumptions and attend closely to how meaning is linguistically realized. Contemporary perspectives emphasize discourse analysis as the examination of language use for communicative purposes within specific social settings, foregrounding the inseparability of linguistic form and context. Related lines of inquiry include genre-based approaches, studies of interpersonal meaning negotiation, and multimodal analyses that account for visual, gestural, and other semiotic resources alongside language.

Building on international developments, scholars within the domestic academic community have advanced distinct yet complementary understandings of discourse analysis. One influential perspective frames it as a method for investigating meaning construction, focusing on how linguistic forms interact with context to generate emotional resonance and layered significance. Another approach, grounded in functional linguistics, treats discourse analysis as an integrative interpretive activity centered on language-in-use. It emphasizes holistic examination of both fragmented and extended linguistic material, with particular attention to cognitive processes underlying discourse production and comprehension—including mental representations and conceptual models. A third line of inquiry prioritizes the interplay between discourse structure and its constituent elements, analyzing how coherence is achieved through structural alignment and functional integration within contextually situated communication.

In summary, discourse analysis is fundamentally a linguistic methodology involving the systematic, empirically grounded examination of language composition to illuminate patterns of use, function, and effect. It entails methodical scrutiny of textual architecture—such as logical progression across sentences and paragraphs, rhetorical patterning, and stylistic configuration—to uncover how meaning is encoded and negotiated [14, 15]. Research consistently affirms the pedagogical value of integrating discourse-analysis strategies with vocabulary development in reading instruction. Explicit teaching of discourse markers—including transitions, reference devices, and substitution patterns—strengthens learners' capacity to parse textual structure and infer implicit relationships. Likewise, targeted vocabulary instruction that addresses semantic networks, morphological families, and collocational behavior reinforces lexical knowledge, thereby supporting deeper comprehension and more accurate interpretation.

3. Objectives of Study

1. Investigate the effectiveness of vocabulary knowledge and discourse analysis skills in enhancing reading comprehension among middle school students in China.
2. Examine the extent to which inferential reasoning and discourse analysis skills influence students' reading comprehension performance, as well as the mediating role of vocabulary knowledge.
3. Assess whether vocabulary knowledge and discourse analysis skills contribute to improvements in students' reading comprehension scores and overall proficiency.

4. Methods of Study

4.1. Research Hypothesis

1. There is no significant relationship between vocabulary knowledge and reading comprehension performance.
2. There is no significant relationship between inferential reasoning ability and reading comprehension performance.
3. There is no significant relationship between discourse analysis skills and reading comprehension performance.

4.2. Participants

A total of 120 middle school students from Dalian City, Liaoning Province, participated in assessments measuring vocabulary knowledge, inferential reasoning, text structure awareness, word-level processing, discourse analysis skills, and reading comprehension.

4.3. Measures

Prior to quantitative analysis, descriptive statistics were applied to the reading comprehension scores. Correlations among the measured dimensions were examined before conducting multiple linear regression analysis. To identify factors influencing reading comprehension performance, a multiple linear regression model was employed with reading comprehension scores as the dependent variable and vocabulary knowledge, inferential reasoning ability, and discourse-analysis skills as independent variables. The model was specified as $Y=\beta1+\beta2X1+\beta3X2+\beta4X3+\mu$, where Y denotes the reading comprehension score, X1, X2, and X3 represent vocabulary, making inferences, and discourse-analysis skills respectively, $\beta2$, $\beta3$, and $\beta4$ are the estimated coefficients, $\beta1$ is the intercept term, and μ is the random error term accounting for unobserved variability.

4.4. Procedures

4.4.1. Text-Structure

The term 'text/discourse' originates from functional linguistics and refers to any oral or written passage of indefinite length, underscoring the inherent indeterminacy of the concept. Some scholars characterize discourse as an abstract theoretical construct, foregrounding the sentence's role in realizing meaning within larger linguistic units. Others define it as a coherent sequence of sentences, emphasizing structural and

functional integration. A widely adopted perspective holds that any passage fulfilling a specific communicative purpose within a given context qualifies as discourse. Within the framework of national curriculum standards, discourse is understood as pedagogical material employed in classroom instruction—including oral, written, and multimodal forms such as diagrams—and is often referred to interchangeably as 'teaching text' or 'teaching discourse'. While definitions vary across scholarly traditions, a consensus emerges around discourse as purposefully deployed language serving instructional aims. Synthesizing these perspectives, discourse may be characterized as a natural language unit composed of successive sentences or paragraphs that establishes thematic focus, situates learning within a meaningful context, and supports communicative functionality and semantic completeness. Given the scope of this study, only written discourse is examined; oral forms are excluded [4, 16]. The central objective is to investigate how English vocabulary knowledge and discourse-analysis skills contribute to reading comprehension among middle school students in China, grounded in the conceptual understanding outlined above.

4.4.2. Making Inference

Students are guided to make inferences about the connotations of key statements by drawing on contextual clues [17, 18]. In intensive reading, context significantly supports the interpretation of vocabulary and the pragmatic meaning of sentences. However, due to developing reading proficiency, middle school students often exhibit limited sensitivity to linguistic nuance and tend to grasp only the literal meaning of sentences that carry rich extra-linguistic implications. Consequently, a central challenge in reading instruction lies in helping students move beyond surface-level comprehension to discern the context-dependent pragmatic meaning of utterances.

4.4.3. Guessing Words through Context

Inferring the meaning of unfamiliar words from surrounding context is a foundational skill for middle school English reading comprehension, and such vocabulary-in-context questions frequently appear in standardized reading assessments. Learners are encouraged to become familiar with established strategies to support accurate interpretation. Developing proficiency in contextual word inference enables students to determine precise meanings based on how words function within sentences and broader textual structures [19, 20]. Vocabulary does not occur in isolation; rather, it connects meaningfully—through semantic relationships, grammatical patterns, and logical coherence—to adjacent words, sentences, and entire passages. Consequently, multiple evidence-based inference techniques exist. In lexical hierarchies, some terms express broad categories that encompass more specific members; these include hypernyms (general terms) and hyponyms (specific instances). Educators may leverage such semantic relationships to scaffold students' inferential reasoning.

For instance, he went to see a physician yesterday and the doctor advised him to give up smoking [21].

The terms 'doctor' and 'physician' refer to the same role, and the prior mention of 'doctor' supports interpreting 'physician' as a medical professional [22, 23].

I know what you think: pizza for breakfast? But the truth is, you can have last night's leftovers in the morning if you want [24].

The context indicates that 'leftovers' functions as a noun meaning food remaining from a previous meal, with 'last night' serving as a temporal cue reinforcing this interpretation [25].

The nine short films to be screened range from a Claymation-produced biography of B. B. King to a sci-fi adventure set in the year 3005.

Given that the passage describes children creating cinematic works, 'shorts' here denotes short films—concise, self-contained motion picture productions [20, 26].

4.4.4. Guessing Words through Contextual Inference Method

Difficulty in determining the meaning of unfamiliar words during English reading remains a primary barrier to students' reading proficiency. Inferring word meaning from contextual clues is both a core skill assessed across standardized English examinations and a key objective in contemporary English language instruction. In high school English reading instruction, deliberate practice in contextual inference enables students to rapidly and accurately grasp word meanings, thereby alleviating vocabulary-related challenges, enhancing overall reading comprehension, improving performance on reading assessments, and fostering greater engagement with English texts. Word meaning inference is inherently dependent on linguistic context; the surrounding textual environment provides essential cues for accurate semantic prediction.

For example, What does the underlined phrase “run out of steam” in paragraph 3 most likely mean? (Tianjin, 2019, Paper 52)

Based on the surrounding text, the sentence containing the underlined phrase indicates that some individuals become exhausted or lose motivation well before reaching life’s midpoint—suggesting diminished enthusiasm or stamina for sustained effort. Thus, 'Loss of interest in continuing studies' is the most appropriate interpretation [27].

4.4.5. Reading Skills-Selecting Main Idea When Reading

Teachers should use instructional procedures to help students determine the importance of what they are reading, or more specifically, the main ideas of the text [28]. Some evidence-based guidelines for teaching students to identify key information in texts are listed below.

Build on students' background knowledge by using concrete examples and connecting new content to prior learning and previous readings.

Organize students' knowledge of reading strategies—for example, by helping them connect existing understanding to upcoming texts through attention to genre, context, main and supporting ideas, topic sentences, and logical connectives [29].

Model the reading process explicitly. Effective modeling techniques include recalling relevant prior knowledge (e.g., 'This reminds me of...') and making reasoned predictions (e.g., 'This may occur because...').

Conduct ongoing reading assessments to support the transfer and application of reading skills across contexts.

Use a variety of formative assessment tools to monitor students’ progress and proficiency in applying reading strategies [30]. Encourage transfer by guiding students to apply learned skills to new but structurally similar texts, thereby strengthening connections between prior and emerging knowledge.

Literal questions focus on explicit, text-based information and prompt recall of details essential for comprehension [31]. For example, 'What did the police officer observe in the victim's room?' is more purposeful than 'What was the police officer's name?'

Inferential questions require students to integrate textual evidence with background knowledge to construct meaning. These include prompts that activate prior understanding (e.g., 'What leads you to that conclusion?') or invite reasoned prediction (e.g., 'Based on current developments, what might happen next to the people stranded on the island?').

Posing both literal and inferential questions before reading helps students clarify their purpose, activate relevant background knowledge, and anticipate the text’s central message [32, 33]. Such questions may be posed by teachers or generated collaboratively by students. Techniques to foster thoughtful questioning include:

1. Identify one or two pivotal moments in the narrative, pause reading, and ask students to infer likely developments.
2. Ask students to substantiate their predictions using specific textual evidence and personal experience.
3. As reading continues, guide students to compare initial predictions with actual outcomes, reflect on discrepancies, and revise interpretations—articulating clear reasoning for any adjustments.

4. Encourage metacognitive reflection through structured reading journals that address both content and language features.
5. Support creative response modes, such as dramatization, visual representation, or constructing storyline flowcharts.

The following Figure 2 displays a model of the information-processing of the reading to instruct skills analysis.

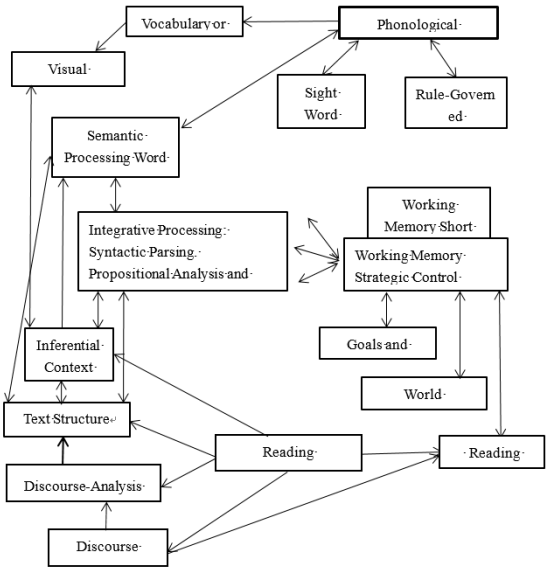

Figure 2. Model of the Information-Processing of the Reading to Instruct Skills Analysis

Vocabulary breadth and discourse analysis skills enhance reading efficiency. A robust vocabulary enables rapid word recognition and comprehension, reducing cognitive load associated with decoding unfamiliar terms. Likewise, discourse analysis skills support efficient skimming, strategic identification of salient information, and accurate extraction of main ideas—thereby improving overall text comprehension speed and depth.

5. Results

This section presents descriptive statistics, correlation analysis, multiple linear regression analyses, and path analyses. The multiple linear regression analyses include the model summary, ANOVA, and model regression coefficients.

5.1. Descriptive Statistics

Descriptive statistics were computed for Vocabulary Knowledge (VK), Making Inference (MI), Discourse-Analysis Skills (DAS), and Reading Comprehension (RC), as presented in Table 1 (As shown in Table 2).

MI: Making Inference
DAS: Discourse-Analysis Skills
RC: Reading Comprehension

Table 1. Descriptive Statistics

	N	Minimum	Maximum	Average	Standard Error	Standard Deviation	ANOVA
VK	120	2	5	3.60	.306	.966	.933
MI	120	2	5	3.60	.400	1.265	1.600
DAS	120	1	5	3.90	.458	1.449	2.100
RC	120	2	5	4.00	.333	1.054	1.111

Table 2. Reliability Check of the Questionnaire

Cronbach's Alpha	N of Items
.912	120

VK: Vocabulary Knowledge

Table 1 reports the minimum, maximum, mean, standard error, standard deviation, and ANOVA results for VK, MI, DAS, and RC.

5.2. Reliability

Reliability refers to the consistency of measurement outcomes. A measurement instrument is deemed reliable when it yields stable and reproducible results across repeated administrations [34]. To ensure reliability, the questionnaire items were rigorously reviewed and subjected to internal consistency analysis. A widely accepted threshold for acceptable reliability is a Cronbach's alpha coefficient of at least 0.70. Using SPSS 29.0 for Windows, the computed Cronbach's alpha for the questionnaire was 0.912—well above the recommended threshold. This indicates high internal consistency, supporting the questionnaire's suitability for use in this study (see Table 2).

5.3. Correlation Analyses

Prior to conducting the multiple linear regression analysis, the correlations among the dimensions were examined, as presented in Table 3.

Correlation is significant at the 0.01 level (two-tailed).

Table 3. Correlation

		VK	MI	DAS	RC
VK	Pearson correlation	1	.491	.603	.764*
	Significance (Double tailed)		.002	.001	.003
MI	Pearson correlation	.491	1	.885**	.833**
	Significance (Double tailed)	.000		.001	.003
DAS	Pearson correlation	.603	.885**	1	.873**
	Significance (Double tailed)	.000	.001		.001
RC	Pearson correlation	.764*	.833**	.873**	1
	Significance (Double tailed)	.000	.003	.001	

As shown in Table 3, the correlation coefficients among vocabulary, inference-making, discourse-analysis skills, and reading comprehension all exhibit significance levels below 0.01. This result rejects the null hypothesis, indicating a significant correlation between reading comprehension scores and vocabulary knowledge, inference-making abilities, and discourse-analysis skills.

5.4. Multiple Linear Regression Analyses

The above analysis indicates a significant correlation between reading comprehension scores and vocabulary knowledge, as well as skills in making inferences and conducting discourse analysis; however, the primary determinants of reading comprehension performance and their relative contributions remain unclear and require further investigation [1, 35].

5.4.1. Model Summary

Model Summary reports the adjusted R-squared value, which reflects the proportion of variance in the dependent variable explained by the independent variables. A higher adjusted R-squared indicates stronger explanatory power. Table 4 presents these metrics.

Table 4 summarizes the regression model. The adjusted R-squared is 0.817, indicating that the explanatory variables collectively account for 81.7% of the variance in the

dependent variable [36]. The significance F-value is 0.004, which is less than the 0.05 threshold, confirming that the overall model is statistically significant and fits the data well.

Table 4. Model Summary

Model	R	R Square	After adjustment R square	Errors in standard estimation	Change Statistics				
					R-squared change amount	F Amount of change	Degree of freedom 1	Degree of freedom 2	Significance F Amount of change
1	.937a	.878	.817	.451	.878	14.391	3	6	.004

a. Predictor variable: (constant), x3, x1, x2

Correlation is significant at the 0.05 level (two-tailed).

5.4.2. Anova

Analysis of Variance (ANOVA) is a statistical method used to test the significance of differences among two or more group means [5, 37]. The results are presented in Table 5. Predicted variables: Constant, VK, MI, DAS

Table 5. Anova

Model		Sum of Squares	Freedom	Mean Square	F	Significance
1	Regression	8.780	3	2.927	14.391	.004b
	Residual	1.220	6	.203		
	Total	10.000	9			

Dependent variable:RC

As shown in Table 5, a p-value greater than 0.05 indicates that the factor exerts minimal influence on the dependent variable and no statistically significant difference exists among groups; conversely, a p-value less than 0.05 suggests a significant effect on the dependent variable and meaningful intergroup differences. For the multivariate linear regression model, the observed F-statistic is 14.391, with an associated p-value effectively zero. At the significance level of 0.004 (< 0.05), the null hypothesis—stating that all regression coefficients equal zero simultaneously—is rejected. This confirms a statistically significant linear relationship between the explanatory variables (Constant, VK, MI, DAS) and the dependent variable RC, supporting the appropriateness of the linear model. Specifically, discourse-analysis skills, inferencing ability, and vocabulary knowledge each exert a significant effect on reading comprehension.

5.4.3. Model Regression Coefficients

The regression coefficient is a parameter in a regression equation that quantifies the magnitude of the influence of an independent variable x on a dependent variable y. A positive regression coefficient indicates that y increases as x increases, whereas a negative regression coefficient indicates that y decreases as x increases. For example, in the regression equation $Y = bX + a$, the slope b represents the regression coefficient, signifying that for each unit change in X, Y changes by an average of b units. Table 6 presents the regression coefficients.

Table 6. Coefficient [A]

Model		Standardized Coefficient	t	significance	95.0% Confidence Interval of B		Correlation			Collinearity Statistics	
		Beta			lower limit	upper limit	zero-order	inclined to one side	part	tolerance	VIF
1	(constant)		0.926	0.355	-1.040	1.976					
	x1	0.285	4.509	0.000	-.047	.914	0.879	0.296	0.12	0.179	5.599
	x2	0.33	5.459	0.000	-.329	.928	0.88	0.351	0.146	0.196	5.105
	x3	0.349	5.729	0.000	-.370	.829	0.883	.366	0.153	0.192	5.203

Dependent Variable: Reading Comprehension

Table 6 displays the coefficient estimates from the regression model. The t-statistic for the constant term is 0.926, with a corresponding p-value of 0.335, exceeding the 0.05 significance threshold; thus, the constant is not statistically significant. In contrast, the t-statistic for vocabulary knowledge is 4.509 ($p < 0.001$), indicating a statistically significant positive effect on reading comprehension performance. Similarly, the t-statistic for making inference is 5.459 ($p < 0.001$), and that for discourse analysis skills is 5.729 ($p < 0.001$), both confirming statistically significant positive contributions to reading comprehension scores.

Variance inflation factors (VIFs) for vocabulary knowledge, making inference, and discourse analysis skills are 5.599, 5.105, and 5.203, respectively—all below 10—indicating no substantial multicollinearity among the predictors. The final regression model is:

$$Y=0.285X1+0.33X2+0.349X3$$

where Y denotes the reading comprehension score, and X_1, X_2, and X_3 represent vocabulary knowledge, making inference ability, and discourse analysis skills, respectively [38, 39]. Specifically, a one-unit increase in vocabulary knowledge corresponds to a 0.285-point increase in reading comprehension score; a one-unit increase in making inference ability corresponds to a 0.33-point increase; and a one-unit increase in discourse analysis skills corresponds to a 0.349-point increase.

5.5. Path Analyses

Path analysis is used to examine linear relationships among multiple independent variables and a dependent variable. Figure 3 illustrates the hypothesized structural relationships among the variables examined in this study.

M: Making Inference
V: Vocabulary Knowledge
T: Text Structure
Dis: Discourse-Analysis Skills
Re: Reading Comprehension

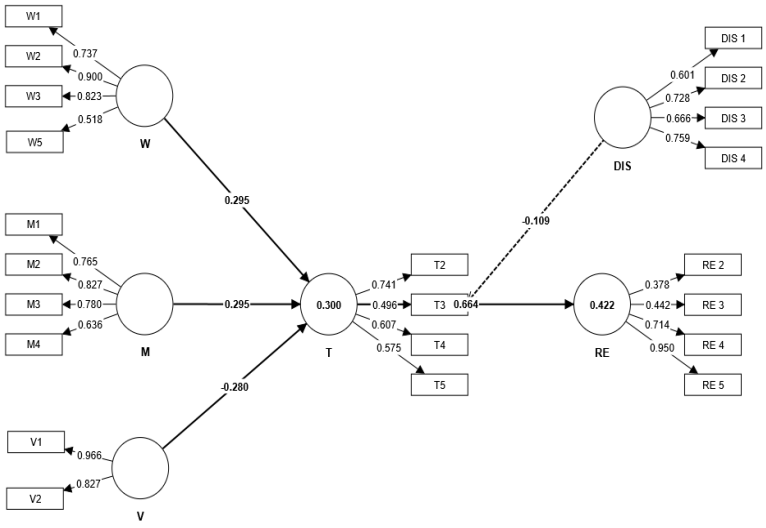

Figure 3. Path Analysis

W: Word

As shown in Figure 3, all direct paths from word knowledge, making inference, vocabulary knowledge, text structure, and discourse-analysis skills to reading comprehension were statistically significant. Additionally, discourse-analysis skills exerted a significant indirect influence on reading comprehension through mediating pathways [40].

5.6. Data Analysis

To provide a more structured analysis of the data obtained in this study on the effectiveness of English vocabulary and discourse-analysis skills in reading comprehension among Chinese middle school students, the following analytical procedures were employed: Descriptive Statistics: Means, standard deviations, and ranges were calculated for English vocabulary scores, discourse-analysis skill scores, and reading comprehension scores to summarize the distribution and central tendency of each variable [41, 42]. Correlations: Pearson correlation coefficients were computed to examine the strength and direction of linear relationships among the three variables, along with associated significance levels. Hypothesis Testing: Appropriate parametric or nonparametric tests were conducted to evaluate predefined hypotheses regarding group differences or associations; test statistics, p-values, and effect sizes were reported to assess statistical and practical significance. Regression Analysis: Multiple linear regression was performed to quantify the predictive contributions of vocabulary and discourse-analysis skills to reading comprehension performance; standardized and unstandardized regression coefficients, confidence intervals, and significance levels were interpreted to clarify the magnitude and direction of each predictor's influence.

6. Discussion

In this thesis, the collected questionnaire data were processed and analyzed using descriptive statistics to examine reading comprehension, followed by correlation analysis to assess the relationships between reading comprehension scores and vocabulary knowledge, making inference ability, and discourse-analysis skills. Building on these analyses, multiple regression models were constructed with reading comprehension scores as the dependent variable and vocabulary knowledge, making inference ability, and discourse-analysis skills as independent variables, yielding the following conclusions:

1. Vocabulary knowledge exhibits a strong positive association with reading comprehension scores; for each unit increase in vocabulary knowledge, reading comprehension scores increase by 0.285.

2. Making inference ability shows a strong positive association with reading comprehension performance; for each unit increase in making inference ability, reading comprehension performance increases by 0.33.
3. Discourse-analysis skills demonstrate a strong positive association with reading comprehension performance; for each unit increase in discourse-analysis skills, reading comprehension performance increases by 0.349.

The findings indicate that robust English vocabulary knowledge and well-developed discourse-analysis skills are positively associated with reading comprehension. These results may inform pedagogical practice and curriculum development in Chinese middle school English education. Furthermore, by aligning the empirical findings with established theoretical perspectives, this study contributes to refining and contextualizing relevant theoretical frameworks.

Implications: Both teachers and students observed benefits from applying the vocabulary- and discourse-analysis–integrated instructional model in junior high school English reading classes [43, 44]. Optimizing the advantages of this model while mitigating its limitations presents an important consideration for middle school English educators.

First, teachers can embed vocabulary instruction and discourse analysis within authentic contexts [29, 45]. Prior explicit instruction on foundational vocabulary concepts and discourse structures helps ensure students develop a sound conceptual understanding. Students should also be guided to analyze texts across diverse genres, as each genre features distinct organizational patterns, linguistic conventions, and communicative purposes. Appropriate methodological choices and content selection can enhance student engagement, initiative, and analytical thinking. Moreover, students have become proficient in applying this instructional approach to their regular English reading practice.

Second, teachers can integrate the vocabulary- and discourse-analysis–integrated model with complementary pedagogical approaches to deepen text comprehension. Designing accessible, interactive classroom activities—such as those aligned with task-based language learning—can strengthen students' functional language use and overall proficiency. Diverse, purposeful tasks help sustain motivation and intellectual curiosity. Additionally, incorporating cultural context related to vocabulary and target-language usage supports the development of intercultural awareness. Alongside improving core reading competencies, broadening students' cultural knowledge and fostering a global outlook remain essential educational goals. While promising outcomes have emerged from implementing this model, certain challenges have also become evident. Sustained refinement is needed to maximize its effectiveness, explore synergies with other evidence-informed teaching strategies, address identified limitations, and better align instruction with learners' evolving needs.

Finally, sufficient vocabulary knowledge and foundational grammatical understanding are essential prerequisites for effective discourse analysis; therefore, systematic attention to language fundamentals must remain integral to reading instruction [1, 33].

7. Conclusions

English vocabulary and discourse analysis skills significantly enhance reading comprehension among Chinese middle school students. Targeted instruction in these areas supports broader English proficiency development and academic achievement. These competencies enable learners to infer the meanings of unfamiliar words or phrases by drawing on contextual clues within a text. Leveraging lexical knowledge alongside awareness of discourse structures allows students to make reasoned semantic approximations, thereby sustaining comprehension despite encountering unknown vocabulary.

This study demonstrates that a vocabulary- and discourse-oriented instructional model yields measurable improvements in middle school English reading: (1) increased student engagement with reading materials; (2) enhanced capacity to identify main ideas

and respond accurately to inference- and prediction-based questions; and (3) improved overall reading performance. The findings confirm that integrating vocabulary and discourse analysis into English reading instruction strengthens discourse awareness, deepens textual understanding, and elevates reading outcomes. The model holds both theoretical significance and practical applicability for English language teaching in junior high school contexts. Relative to prior work, this study contributes three key innovations: first, its exclusive focus on junior high school English reading; second, the systematic integration of lexical and discourse-level strategies within a single pedagogical framework; and third, empirical validation of the model's efficacy through classroom-based implementation and outcome assessment.

7.1. Improvement of Students' Reading Ability

Reading ability is a critical component of overall English proficiency. When engaging with texts, students often overlook structural and functional features, limiting their capacity to interpret texts holistically. A limited repertoire of reading strategies further constrains skill development. Enhancing student engagement in English learning supports greater motivation and self-directed learning behaviors, thereby improving learning efficiency and outcomes. To address these challenges, instructional approaches emphasizing vocabulary and discourse analysis have been integrated into middle school English reading instruction. This model strengthens discourse awareness while concurrently advancing reading comprehension and analytical capacity.

7.2. Teachers' Teaching Skills Development

To effectively apply vocabulary and discourse analysis in English reading instruction, teachers are encouraged to engage extensively with scholarly and pedagogical literature. Mastery of discourse analysis principles has enhanced professional competence and instructional quality, yielding mutual benefits for educators and learners while supporting broader adoption of innovative teaching approaches in junior high school English classrooms.

7.3. Impact on the Traditional Reading Teaching Mode

The traditional approach to English reading instruction in junior high school emphasizes vocabulary, grammar, and discrete language points within texts, which limits the development of holistic reading competence. Effective reading comprehension aims not only at decoding linguistic elements but also at identifying the text's central ideas, organizational structure, and authorial purpose. Integrating explicit instruction in reading strategies supports learners in efficiently extracting key information and cultivating sustainable, self-regulated reading habits. Furthermore, contemporary curriculum frameworks emphasize that language learning must extend beyond linguistic knowledge to include meaningful engagement with diverse cultural contexts—particularly those associated with English-speaking communities—thereby fostering intercultural awareness and global perspective.

Given these considerations, incorporating vocabulary and discourse analysis into junior high school English reading instruction is both timely and pedagogically necessary. This shift actively transforms conventional teaching practices. The discourse analysis model offers mutual benefits: it enhances student motivation, enlivens classroom interaction, strengthens analytical and critical thinking capacities, and increases instructional efficiency.

The advantages of strengthening vocabulary and discourse analysis skills extend well beyond secondary education. Learners equipped with robust reading comprehension strategies are better prepared for academic success in higher education and professional settings. As global communication intensifies, proficiency in English—including nuanced textual interpretation—becomes increasingly vital for navigating complex, cross-cultural environments. Establishing a rigorous foundation in these competencies during adolescence empowers students to engage confidently with global challenges.

References

1. R. C. Anderson and P. D. Pearson, "A schema-theoretic view of basic processes in reading comprehension," Center for the Study of Reading Technical Report, no. 306, 1984.
2. A. Bonyadi, "Discourse analysis and language pedagogy: A review," *Journal of Teacher Education for Sustainability*, vol. 21, no. 1, pp. 128–136, 2019.
3. C. Brown and M. E. Payne, "Five essential steps of processes in vocabulary learning," presented at the TESOL Convention, Baltimore, MD, Mar. 1994.
4. D. H. Karrfalt, "The generation of paragraphs and larger units," *College Composition & Communication*, vol. 19, no. 3, pp. 211–217, 1968.
5. R. P. Carver, "Reading rate: Theory, research, and practical implications," *Journal of Reading*, vol. 36, no. 2, pp. 84–95, 1992.
6. J. W. Creswell and J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage Publications, 2017.
7. R. F. DeVellis, *Scale Development: Theory and Applications*, 2nd ed., vol. 26. Thousand Oaks, CA: Sage Publications, 2012.
8. N. K. Duke and P. D. Pearson, "Effective practices for developing reading comprehension," *Journal of Education*, vol. 189, no. 1–2, pp. 107–122, 2009.
9. J. Firbas, *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge, UK: Cambridge University Press, 1992.
10. M. Foucault, *Madness and Civilization*. London, UK: Routledge, 2003.
11. A. Georgakopoulou, *Discourse Analysis: An Introduction*. Edinburgh, UK: Edinburgh University Press, 2019.
12. T. Givón and H. A. Whitaker, *On Understanding Grammar*, vol. 379. New York, NY: Academic Press, 1979.
13. J. T. Guthrie and N. M. Humenick, "Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement," 2004.
14. M. A. K. Halliday and R. Hasan, *Cohesion in English*. London, UK: Routledge, 2014.
15. J. E. Grimes, *The Thread of Discourse*, vol. 207. Berlin, Germany: Walter de Gruyter, 1975.
16. R. Gill, "Discourse analysis," in *Qualitative Researching with Text, Image and Sound*, vol. 1, 2000, pp. 172–190.
17. J. Harmer, *The Practice of English Language Teaching*. London, UK/New York, NY: Longman, 2001.
18. W. Kintsch, "The role of knowledge in discourse comprehension: A construction-integration model," *Psychological Review*, vol. 95, no. 2, p. 163, 1988.
19. W. Labov and J. Waletzky, "Narrative analysis: Oral versions of personal experience," 1997.
20. B. Laufer, "The lexical profile of second language writing: Does it change over time?," *RELC Journal*, vol. 25, no. 2, pp. 21–33, 1994.
21. B. Laufer and Z. Goldstein, "Testing vocabulary knowledge: Size, strength, and computer adaptiveness," *Language Learning*, vol. 54, no. 3, pp. 399–436, 2004.
22. C. N. Li and S. A. Thompson, "Subject and topic: A new typology of language," in *Subject and Topic*, London, UK: Routledge, 1976, pp. 457–489.
23. R. E. Longacre, *An Anatomy of Speech Notions*, vol. 3. Berlin, Germany: Walter de Gruyter GmbH & Co KG, 2020.
24. W. Mann, "Rhetorical structure theory: Toward a functional theory of text organization," *Text*, vol. 8, no. 3, pp. 167–182, 1987.
25. W. C. Mann, C. M. I. M. Matthiessen, and S. A. Thompson, "Rhetorical structure theory and text analysis," in *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*, 1992, pp. 39–78.
26. P. Meara, "Towards a new approach to modelling vocabulary acquisition," in *Vocabulary: Description, Acquisition and Pedagogy*, 1998, p. 109.
27. K. L. Pike, *Language in Relation to a Unified Theory of the Structure of Human Behavior*, vol. 24. Berlin, Germany: Walter de Gruyter GmbH & Co KG, 2015.
28. I. Parker, *Discourse Dynamics: Critical Analysis for Social and Individual Psychology*. London, UK: Routledge, 2014.
29. M. McCarthy and R. Carter, *Language as Discourse: Perspectives for Language Teaching*. London, UK: Routledge, 2014.
30. M. Pressley and P. Afflerbach, *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. London, UK: Routledge, 2012.
31. L. Polanyi, "A formal model of the structure of discourse," *Journal of Pragmatics*, vol. 12, no. 5–6, pp. 601–638, 1988.
32. H. Sacks and G. Jefferson, *Lectures on Conversation*. Dordrecht, Netherlands: Springer, 1995.
33. E. A. Schegloff and H. Sacks, "Opening up closings," 1973.
34. H. Sacks, E. A. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation," *Language*, vol. 50, no. 4, pp. 696–735, 1974.
35. S. Blum-Kulka, "Discourse pragmatics," in *Discourse as Social Interaction*, vol. 2, 1997, pp. 38–63.
36. S. Webb, "Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge," *Studies in Second Language Acquisition*, vol. 27, no. 1, pp. 33–52, 2005.
37. H. G. Widdowson, "An applied linguistic approach to discourse analysis," Ph.D. dissertation, Univ. of Edinburgh, Edinburgh, UK, 1973.

38. S. Webb, "Receptive and productive vocabulary sizes of L2 learners," *Studies in Second Language Acquisition*, vol. 30, no. 1, pp. 79–95, 2008.
39. J. Woodilla, "Workplace conversations: The text of organizing," in *Discourse and Organization*, 1998, pp. 31–50.
40. L. Li, D. Zhu, and X. Wu, "The effects of vocabulary breadth and depth on reading comprehension in middle childhood: The mediator role of listening comprehension," *Reading & Writing Quarterly*, vol. 37, no. 4, pp. 336–347, 2021.
41. J. Yan and J. Jiang, "Discourse analysis and High School English Reading Teaching," *International Journal of Educational Development*, vol. 1, no. 2, pp. 72–79, 2025.
42. Y. Dong, J. Hu, X. Wu, H. Zheng, and X. Peng, "The evidence of different learning environment learning effects on vocabulary size and reading comprehension," *Frontiers in Psychology*, vol. 9, p. 1914, 2018.
43. H. Zhong, "Multidimensional vocabulary knowledge: Development from receptive to productive use," in *Current Perspectives in Second Language Vocabulary Research*, 2012, pp. 23–55.
44. S. Zhou, "Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners," *Asian Social Science*, vol. 6, no. 10, p. 14, 2010.
45. G. Youmans, "A new tool for discourse analysis: The vocabulary-management profile," *Language*, vol. 67, no. 4, pp. 763–789, 1991.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.