Theme-Based, Sheltered, or Adjunct? Evaluating CBI Models for Improving English Reading Skills in Chinese EFL Classrooms

Lu Zhang[1], Hanita Hanim Ismail[1], Nur Ainil Sulaiman[1]

[1] Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Correspondence: Hanita Hanim Ismail, Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia. E-mail: hanitahanim@ukm.edu.my

Received: August 13, 2025 Accepted: September 22, 2025 Online Published: October 21, 2025

doi:10.5430/wjel.v16n2p282 URL: https://doi.org/10.5430/wjel.v16n2p282

Abstract

Content-Based Instruction (CBI) has been widely promoted as an effective approach for integrating language development with subject matter learning in English as a Foreign Language contexts. However, limited empirical evidence exists comparing the effects of different CBI models on students' reading proficiency, particularly within Chinese tertiary education. Existing studies often focus on general outcomes or perceptions, leaving a gap in understanding how specific instructional models influence distinct reading sub-skills. This study aimed to address this gap by comparing the effectiveness of three CBI models (theme-based, adjunct, and sheltered) on university students' English reading proficiency, focusing on three sub-skills: understanding explicit information, understanding implicit information, and using linguistic features to understand texts. Employing a mixed-methods design, the study involved 105 Chinese university students across three intact classes, with quantitative data collected from the English reading pretest and posttest and qualitative data collected from classroom observations and semi-structured interviews. Quantitative results indicated that while all groups improved significantly from pretest to posttest, the theme-based model led to significantly higher gains in total reading scores and in the ability to use linguistic cues. Qualitative findings suggested that the instructional design, input organization, and opportunities for textual scaffolding varied across models and contributed to the outcomes. These findings offer practical implications for EFL curriculum planners and educators seeking to align reading instruction with model-specific strengths for more effective literacy development.

Keywords: CBI, theme-based CBI, sheltered CBI, adjunct CBI, English reading

1. Introduction

In the era of globalization and international academic exchange, reading proficiency in English has become a critical skill for university students, particularly in English as a Foreign Language (EFL) contexts. Reading is not only foundational for academic success but also essential for lifelong learning, as it enables learners to access, interpret, and engage with a wide range of English-language texts (Grabe & Yamashita, 2022). In many EFL environments, however, reading instruction remains confined to traditional grammar-translation or test-oriented approaches, which often limit students' ability to interact meaningfully with content (Komarawan, 2024). This situation is especially pressing in China's higher education system (Zheng et al., 2025), where the demand for academically competent English readers continues to rise due to internationalization policies and the prevalence of English-medium academic materials.

To address these challenges, Content-Based Instruction (CBI) has emerged as a promising pedagogical framework. CBI integrates language learning with subject content and offers learners a more meaningful, contextualized, and cognitively engaging experience (Cheng, 2024; Suvonova, 2023). The implementation is underpinned by different models (Brinton & Snow, 2017; Snow & Brinton, 2023). For example, the theme-based language instruction model is a language course organized around topical themes where content serves as the vehicle for achieving language objectives, emphasizing topics of interest to build vocabulary and comprehension. The sheltered content instruction model is a content course adapted for second-language learners by using simplified English, explicit vocabulary and discourse scaffolds, so students develop language while mastering academic subject matter. The adjunct model links a language course and a discipline course taken concurrently with coordinated syllabi, so learners develop targeted language skills in parallel with the concepts, genres, and tasks of the content class. While CBI has been widely endorsed for its theoretical merits and positive outcomes in general, comparative evidence on the effectiveness of these three CBI models is limited, as previously called for by Snow and Brinton (2023).

Moreover, few studies have focused specifically on reading proficiency as a target outcome of CBI, even though reading remains a cornerstone of academic language use and a key predictor of success in higher education. This neglect is surprising given that reading in a second or foreign language (L2) involves a wide range of complex and layered cognitive processes (Birch, 2014; Grabe & Yamashita, 2022), including locating and synthesizing explicit information, inferring implied meanings, and recognizing and interpreting linguistic cues such as cohesion devices and syntactic structures (Sadeghi, 2021). These sub-skills are particularly important in university settings, where students are expected to engage with discipline-specific texts that demand high levels of comprehension, analytical thinking, and

independent interpretation (Amin et al., 2021; Masharipova, 2022). However, most existing studies on CBI either treat reading as a secondary skill, focusing instead on speaking (Hu & AlSaqqaf, 2024), writing (e.g. Chekol et al., 2023; Sariani et al., 2022), vocabulary (e.g. S. Li, 2020), grammar (e.g. Ponna et al., 2024), or general language development (e.g. Hu et al., 2022; Suvonova, 2023). Also, these studies often fail to distinguish among the constituent elements that make up reading comprehension (e.g. Fikni et al., 2024; Masharipova, 2022). As a result, there remains a lack of clarity about which reading sub-skills are most effectively supported by different CBI models, as noted in previous systematic reviews of CBI (Zhang et al., 2024; Zhang & Tang, 2024), and how instructional design might be tailored to improve them in EFL university contexts. A closer examination of these sub-skills, and how they respond to specific pedagogical interventions, is therefore crucial for deepening our understanding of CBI's potential in developing academic reading proficiency.

To address these gaps, the present study adopts a mixed-methods design to examine the comparative effectiveness of the three main CBI models (theme-based, sheltered, and adjunct) on Chinese university EFL learners' English reading proficiency characterized by the ability to understand explicit information, understand implied information, and use linguistic features to comprehend texts, as outlined by both scholars (Birch, 2014; Grabe & Yamashita, 2022; Nation, 2009) and English language education (ELE) authorities in China (Ministry of Education of the People's Republic of China & National Language Commission of the People's Republic of China, 2018). Specifically, the study seeks to answer the following research questions:

- What is the effect of each CBI model on Chinese EFL learners' English reading proficiency?
- What are learners' perceptions of their reading development under each CBI model?
- What contextual factors influence the implementation and effectiveness of the three models in the higher education EFL context?

By integrating both quantitative and qualitative data, this study aims to provide a nuanced understanding of how different CBI approaches shape English reading outcomes in Chinese universities. The findings will contribute to the broader field of EFL pedagogy by offering practical insights for educators, curriculum designers, and policymakers aiming to enhance academic reading instruction through content-integrated approaches.

2. Literature Review

2.1 Content-Based Instruction

CBI is an approach to second and foreign language teaching that integrates language instruction with the learning of subject matter. It is grounded in the principle that language is best acquired when it is used as a medium to engage with meaningful content, rather than as an isolated object of study (Brinton & Snow, 2017). Rooted in theories of communicative competence and cognitive psychology, CBI is premised on the idea that language learning is most effective when it occurs in authentic, contextualized, and cognitively demanding environments (Lyster & Ballinger, 2011). Learners acquire language as they use it to understand, process, and discuss subject-specific information, which promotes both linguistic accuracy and academic literacy.

CBI is typically implemented through three main instructional models: theme-based, sheltered, and adjunct instruction (Brinton & Snow, 2017; Snow & Brinton, 2023). In theme-based CBI, language classes revolve around specific content themes, allowing learners to build vocabulary and skills around familiar topics while maintaining language objectives (Suvonova, 2023). Sheltered instruction, more common in mainstream education, involves teaching subject matter in the target language to learners who are not yet proficient, with teachers modifying content delivery to match learners' linguistic needs (Echevarria & Graves, 2014). Adjunct instruction, by contrast, pairs a language course with a content course, and students are expected to apply their language learning directly to the understanding of academic subject matter (Karimi & Ghorbanchian, 2022). Each model reflects varying degrees of integration between language and content, and their effectiveness often depends on contextual factors such as learners' proficiency, curricular goals, and teacher expertise.

CBI has been associated with a range of pedagogical benefits. These include increased learner motivation, enhanced retention of vocabulary and grammar, and improved cognitive engagement (Suvonova, 2023; Zhang et al., 2024). Of particular interest is its impact on L2 reading development, which has received growing attention in recent years. The integration of content-rich texts and cognitively demanding tasks in CBI contexts encourages learners to engage with reading for meaning rather than form, thereby fostering deeper processing strategies (Snow & Brinton, 2023). Studies suggest that CBI can enhance learners' abilities to identify explicit information, infer implicit meaning, and draw on textual features for comprehension, which are all key components of academic reading (Amin et al., 2021; Masharipova, 2022). Moreover, sustained exposure to disciplinary discourse in CBI environments supports the development of genre awareness and critical reading skills (Zhang & Tang, 2024), which are essential for success in academic and professional domains.

In the context of ELE in China, CBI has gained increasing policy and institutional support, particularly within higher education. For example, following the *College English Teaching Guidelines* (College Foreign Language Teaching Steering Committee of the Ministry of Education, 2020), universities have been encouraged to move beyond traditional grammar-translation and exam-oriented approaches by integrating language instruction with disciplinary knowledge. The guidelines explicitly call for the development of English curricula that embed authentic content and cultivate students' capacity to use English in academic and professional settings (Xiong & Jiang, 2020). Similarly, the Double First-Class Initiative (*Shuang Yi Liu*) has incentivized key universities to internationalize their curricula, including through the adoption of content-integrated English courses and bilingual instruction in non-language majors (Y. Li, 2020; Xiong & Jiang, 2020). In practice, these policies have prompted an expansion of theme-based English courses on global issues, cross-cultural

communication, and professional fields such as business (Yu et al., 2024), engineering (Li & Fu, 2021), and medicine (Hu et al., 2025). However, more structurally integrated forms of CBI, such as sheltered and adjunct instruction, are rarely implemented at scale (Hu & AlSaqqaf, 2024). Institutional barriers, including rigid curricular structures, limited teacher training in content-language integration, and concerns about students' English proficiency (Cheng, 2024; Zhang et al., 2024), have hindered deeper adoption.

Despite growing interest in CBI, empirical research in China remains limited in both scope and focus. Most studies have centered on learners' attitudes or broad language gains, often treating CBI as a monolithic approach rather than disaggregating it into distinct models (Zhang et al., 2024; Zhang & Tang, 2024). This has led to a lack of nuanced understanding of how different CBI models (theme-based, sheltered, and adjunct) differentially affect discrete language skills, particularly reading comprehension, which is central to academic success. Furthermore, very few studies have operationalized reading proficiency in terms of its subcomponents, such as understanding explicit information, inferring implied meanings, and interpreting linguistic cues in texts, as emphasized in both international research (Grabe & Yamashita, 2022) and Chinese national assessment standards. As such, there is a pressing need for comparative studies that evaluate the effectiveness of distinct CBI models on L2 reading outcomes in authentic instructional settings. Addressing this gap, the present study adopts a mixed-methods design to examine how theme-based, sheltered, and adjunct CBI influence Chinese university students' English reading proficiency, thereby offering empirical insights into instructional design and policy implementation within the evolving landscape of English language education in China.

2.2 L2 Reading

L2 reading is a cognitively demanding process that encompasses multiple interrelated subskills, including vocabulary knowledge, grammatical understanding, reading fluency, and discourse-level comprehension (Birch, 2014). While all of these are essential for holistic reading development, this study narrows its focus to three specific subskills: the ability to understand explicit information, infer implied meanings, and utilize linguistic features such as cohesion markers and syntactic cues to construct meaning from texts. These subskills were selected not only because they form the core of academic reading as defined by both L2 reading research scholars (Grabe & Yamashita, 2022; Nation, 2009) and the educational policy in the research context, namely China (Ministry of Education of the People's Republic of China & National Language Commission of the People's Republic of China, 2018), but also because they are directly aligned with the kinds of reading tasks students face in university coursework. Unlike vocabulary or grammar, which are often addressed through discrete exercises in traditional language instruction (Mohaideen et al., 2020), these higher-order comprehension skills require contextualized, content-rich input and scaffolded opportunities for inference and synthesis (Suvonova, 2023), which are conditions that are particularly well-supported by CBI.

Within CBI, the three models align with these subskills in distinct ways. Theme-based instruction, organized around familiar topics, supplies coherent lexical fields and repeated discourse patterns that facilitate accurate retrieval of explicit information and strengthen sensitivity to linguistic features such as cohesion markers and clause relations through strategy instruction embedded in topical reading cycles (Suvonova, 2023). Sheltered instruction, which delivers subject matter with modified input and targeted scaffolds, is well suited to building implicit-meaning inferences because teachers can calibrate text difficulty, front-load disciplinary schemata, and make inferential moves explicit while also pausing to draw attention to syntax and connectors when they impede comprehension (Echevarria & Graves, 2014). Adjunct instruction, pairing a language course with a co-enrolled content course, most directly rehearses the integrative processing required in university reading: students apply language work to authentic disciplinary texts, coordinating explicit detail tracking with implicit stance and purpose detection, and leveraging linguistic features to synthesize across sources and genres (Karimi & Ghorbanchian, 2022). Together, these alignments clarify why CBI, by design, creates the contextualized, content-rich conditions that higher-order L2 reading processes demand (Grabe & Yamashita, 2022; Nation, 2009).

Despite the theoretical fit between CBI and reading development, empirical research examining CBI's effects on L2 reading remains limited and uneven. Most existing studies tend to report general language gains or focus on speaking and writing outcomes, often overlooking reading as a discrete skill (Amin et al., 2021; Fikni et al., 2024; Masharipova, 2022). Where reading is addressed, the instructional models of CBI are often not clearly differentiated, making it difficult to draw conclusions about which model is most effective. Nevertheless, some studies offer promising insights. Theme-based CBI, which integrates language instruction around broad, engaging topics, has been found to improve learners' global comprehension, reading motivation, and vocabulary retention (Fikni et al., 2024; Khusniyah & Wadi, 2020; Komarawan, 2024). Sheltered instruction, typically used to deliver subject content in simplified English, has shown potential to support reading fluency and inferencing in content-heavy disciplines such as science and business (Chandler, 2020). Adjunct models, where language and content courses are aligned, are less frequently studied but have demonstrated positive outcomes in contexts where academic reading is a core requirement, helping learners transfer reading strategies to authentic materials (Karimi & Ghorbanchian, 2022).

However, few comparative studies have systematically evaluated how these distinct CBI models influence specific L2 reading subskills. As Zhang and Tang (2024) note in their systematic review, much of the literature treats reading as a homogeneous construct, such as program evaluations that report only a single pre–post reading score, CBI interventions that judge "reading improvement" via a holistic rubric without subskill breakdown, and assessment studies that collapse item performance into one total (e.g., test scores) rather than separating explicit, inferential, and cohesion/syntax-based processes. This practice limits our understanding of how reading instruction can be tailored to meet academic demands in EFL university contexts, particularly in countries such as China, where internationalization pressures require students to access increasingly complex English-language academic texts (Y. Li, 2020). Moreover, the lack of operational clarity in previous studies regarding what constitutes "reading proficiency" and how it is assessed further complicates the picture (Zhang & Tang, 2024).

Addressing this gap, the present study adopts a targeted and skill-specific approach to reading, operationalizing it through the three core subskills of identifying explicit content, inferring meaning, and navigating textual features. By comparing how theme-based, sheltered, and adjunct CBI models affect these subskills, the study contributes to a more nuanced understanding of instructional effectiveness in content-integrated reading pedagogy.

3. Methodology

3.1 Research Context and Design

This study was conducted at a prestigious research-centered university in eastern China that has been actively promoting English language innovation under the national "Double First-Class" initiative. English reading is a core component of the university's EFL curriculum, and ongoing curricular reforms have encouraged the adoption of integrated content-language pedagogies to enhance students' academic literacy. Within this context, the present study employed a multilevel mixed-methods design (Headley & Clark, 2019) to examine the effectiveness of three CBI models (theme-based, sheltered, and adjunct) on students' English reading proficiency.

All three sections used the same prescribed English reading textbook, with contact hours and assessments held constant; what varied was how the materials were framed and scaffolded by model. In the theme-based section, lessons were organized around weekly topics drawn from the textbook, with language objectives prioritized through vocabulary building, strategy instruction, and thematic discussion. In the sheltered section, the textbook readings were treated as subject-matter input, with the instructor modifying delivery through simplified explanations, glosses, and guided practice to make content accessible to learners with developing proficiency. In the adjunct section, the reading course was explicitly linked to a co-enrolled content survey; the same textbook units were synchronized with disciplinary topics, and students were expected to apply their language learning directly to interpreting and synthesizing content-related texts. Workload and homework were kept comparable across groups.

In the quantitative phase of the study, three intact English reading classes were purposively selected based on the instructional model adopted in each. Each class received a distinct type of CBI over the course of one semester, implemented within a reading unit designed to improve academic comprehension. The instructional interventions were designed in collaboration with course instructors to ensure consistency in learning objectives and assessment standards, while varying only in the structure and integration of content and language. Pre- and post-tests measuring students' reading proficiency, which was operationalized through understanding explicit information, inferring implied meanings, and interpreting linguistic features, were administered to assess learning outcomes.

The qualitative phase complemented the quantitative findings by offering deeper insights into classroom processes and learner experiences. During the instructional period, classroom observations were conducted to capture interactional patterns, scaffolding strategies, and learner engagement across the three CBI models. Observation notes focused specifically on how instructors facilitated the development of the three targeted reading subskills. Upon completion of the instructional unit, semi-structured interviews were conducted with selected students and instructors from each group to explore their perceptions of reading development, instructional support, and contextual challenges. This multilevel design allowed for a comprehensive evaluation (Headley & Clark, 2019) of the pedagogical effectiveness and practical viability of the three CBI models in fostering academic reading skills in an EFL university context.

3.2 Participants

Participants in the study were 105 second-year undergraduate students enrolled in a compulsory English reading course at a prestigious research-intensive university in eastern China. A purposive sampling method was employed in the quantitative phase to select three intact classes, each comprising 35 students with informed consent, based on their similar academic backgrounds, English proficiency levels, and program structures. The key sampling criteria included: (1) enrolment in the same academic year, (2) comparable performance on the university's internal English placement test (equivalent to CEFR B1–B2), and (3) availability for the entire duration of the instructional intervention. Each class was assigned to receive a different type of CBI instruction (theme-based, sheltered, or adjunct) within the same 10-week English reading unit.

The participants ranged in age from 19 to 21 years, with a mean age of 19.8 years. Of the 105 students, 61 were female (58%) and 44 were male (42%). They came from a variety of non-English majors, including disciplines such as engineering, business, and social sciences. To minimize instructional bias, each class was taught by a different teacher who had received specialized training in the assigned CBI model. All instructors possessed postgraduate degrees in ELE or applied linguistics and had at least five years of teaching experience in university-level EFL programs.

In the qualitative phase, voluntary sampling was used to recruit 25 students for post-intervention interviews (8 theme-based, 8 sheltered, 9 adjunct). This approach was appropriate because interviews required sustained reflection and disclosure; inviting volunteers ethically minimized coercion in a teacher–student context and increased the likelihood of information-rich accounts from participants motivated to articulate their experiences. To reduce self-selection bias, invitations were extended to all students with balanced caps per section, and volunteers were purposefully varied by gender, major, and proficiency band to broaden perspectives. The final sample size was guided by saturation, reached when additional interviews no longer yielded novel themes (Hennink & Kaiser, 2020). In addition, the three instructors were also interviewed to provide instructional insights and contextual understanding.

3.3 Data Collection and Analysis

To assess learners' English reading proficiency, a standardized English reading test was administered both before and after the intervention.

The test was adopted from authentic College English Test materials (Huang et al., 2022) and consisted of three reading tasks containing a total of 30 multiple-choice items (1 mark each). Because the College English Test reading section is officially designed to assess students' ability to understand explicit and implicit information and to use linguistic features to interpret texts (Ministry of Education of the People's Republic of China & National Language Commission of the People's Republic of China, 2018), it was deemed appropriate for this study. Following expert review, 11 items were aligned with understanding explicit information (e.g., locating factual details), 11 with understanding implied information (e.g., inference, author's intention), and 8 with using linguistic features to understand texts (e.g., referents, connectives, syntactic parsing). Five English assessment specialists independently conducted the item-skill classification using a reading skill taxonomy. Inter-rater reliability was high, with Cohen's kappa = 0.88, and disagreements were resolved through consensus discussion. This process ensured content relevance and construct alignment of the test items with the intended reading sub-skills. The test's face validity was confirmed through alignment with national English proficiency benchmarks. To assess reliability, a test–retest procedure was conducted with 63 non-participating students from a comparable cohort. The correlation between the two administrations over a two-week interval was $r = 0.81$ ($p < .001$), indicating satisfactory stability.

Due to violations of assumptions for parametric testing (non-normal score distributions confirmed via Shapiro–Wilk tests), non-parametric analyses were conducted using Statistical Package for the Social Sciences 28.0. To examine differences in reading proficiency across the three CBI groups, Kruskal-Wallis tests were conducted on pretest and posttest scores separately. Where significant differences were found, Mann–Whitney U tests were used for pairwise comparisons. To explore within-group changes over time, Wilcoxon signed-rank tests were employed to compare pretest and posttest scores for each group to determine whether particular CBI models had differential effects on reading proficiency.

Qualitative data were collected through two main methods: non-participant classroom observations and individual semi-structured interviews. Observations were conducted by the researchers during the intervention phase using an adapted observation protocol based on Hu and Zhang (2024), focusing on teacher scaffolding strategies, learner engagement with content, and visible reading comprehension processes. Observation notes were recorded systematically and later transcribed. Following the intervention, selected students from the three CBI groups were interviewed individually in a semi-structured way. The interviews were conducted in Chinese to ensure clarity and comfort. Questions included prompts such as: "What parts of the reading class helped you most in understanding texts?" and "Can you describe a time when the class helped you read between the lines or figure out what the writer meant indirectly?" Interviews lasted 20–30 minutes and were audio-recorded and transcribed verbatim.

Thematic analysis was employed to analyze the qualitative data, following Braun and Clarke's (2021) six-step approach: familiarization, coding, theme development, review, definition, and reporting. A hybrid approach was adopted, combining inductive coding with deductive themes drawn from the three targeted reading sub-skills. Coding was conducted using NVivo 14.0. Data saturation was reached when no new themes emerged in later interviews. Triangulation across observation and interview data enhanced the credibility and depth of interpretation.

4. Findings

4.1 Quantitative Findings

The descriptive statistics of the English reading test are presented in Table 1. Table 2 presents the results of the Kruskal–Wallis H tests, which were used to examine group differences due to the non-normal distribution of scores, as identified by the Shapiro–Wilk test. At the pretest stage, no statistically significant differences were found among the three instructional groups (theme-based, adjunct, and sheltered) for any of the measured reading sub-skills or the total test score. Specifically, group differences were not significant for understanding explicit information, $\chi^2(2) = 3.21$, $p = .361$; understanding implied information, $\chi^2(2) = 0.51$, $p = .916$; using linguistic features to understand texts, $\chi^2(2) = 2.59$, $p = .460$; or total reading scores, $\chi^2(2) = 1.51$, $p = .681$. These results confirm baseline comparability among the groups prior to the intervention. However, significant differences emerged in the posttest. While no significant group difference was observed for understanding explicit information, $\chi^2(2) = 0.43$, $p = .934$, the groups differed significantly in understanding implied information, $\chi^2(2) = 7.93$, $p = .048$; using linguistic features to understand texts, $\chi^2(2) = 25.59$, $p < .001$; and total reading test scores, $\chi^2(2) = 25.85$, $p < .001$.

Table 1. Descriptive Statistics of English Reading Test

Reading Skill	Class	Pretest			Posttest		
		Mean	Std. Deviation	Std. Error	Mean	Std. Deviation	Std. Error
Understanding Explicit Information	Theme-Based Model	6.20	1.746	0.295	8.26	2.020	0.341
	Adjunct Model	6.54	2.091	0.353	8.09	1.821	0.308
	Sheltered Model	6.77	2.579	0.436	7.97	2.538	0.429
Understanding Implicit Information	Theme-Based Model	6.03	1.823	0.308	8.26	2.119	0.358
	Adjunct Model	6.23	1.942	0.328	7.94	1.282	0.217
	Sheltered Model	6.31	2.518	0.426	7.91	1.931	0.326
Using Linguistic Features to	Theme-Based Model	5.26	1.094	0.185	7.23	1.140	0.193

Understand Texts	Adjunct Model	5.06	1.327	0.224	6.17	1.200	0.203
	Sheltered Model	5.29	1.250	0.211	6.17	1.485	0.251
Total Reading Test Score	Theme-Based Model	17.49	2.454	0.415	23.74	2.769	0.468
	Adjunct Model	17.83	3.451	0.583	22.20	2.541	0.430
	Sheltered Model	18.37	4.124	0.697	22.06	3.702	0.626

Table 2. Kruskal-Wallis Test Statistics of English Reading Test

Test	Test Statistics[a,b]	Reading Skill			
		Understanding Explicit Information	Understanding Implicit Information	Using Linguistic Features to Understand Texts	Total Reading Test Score
Pretest	Chi-Square	3.208	0.512	2.588	1.505
	Asymp. Sig.	0.361	0.916	0.460	0.681
Posttest	Chi-Square	0.430	7.928	25.592	25.850
	Asymp. Sig.	0.934	0.048	0.000	0.000

a. Kruskal-Wallis Test

b. Grouping Variable: Class

To further explore the group differences identified in the posttest phase, Mann–Whitney U tests were conducted for pairwise comparisons among the three instructional models (see Table 3). For the sub-skill "Understanding Implicit Information," although the Kruskal–Wallis test yielded a statistically significant group-level effect (p = .048), none of the pairwise comparisons reached significance. The mean posttest scores were comparable across the theme-based model (M = 8.26, SD = 2.12), adjunct model (M = 7.94, SD = 1.28), and sheltered model (M = 7.91, SD = 1.93). Mann–Whitney U values ranged from 559.00 to 590.50, with p-values between .523 and .792, indicating no reliable differences between instructional models in fostering students' ability to infer implied information. In contrast, for the sub-skill "Using Linguistic Features to Understand Texts," the theme-based model (M = 7.23, SD = 1.14) significantly outperformed both the adjunct model (M = 6.17, SD = 1.20; U = 315.50, Z = −3.63, p < .001) and the sheltered model (M = 6.17, SD = 1.49; U = 367.00, Z = −3.06, p = .002). However, no significant difference was observed between the adjunct and sheltered groups (U = 607.50, Z = −0.06, p = .952), suggesting that only the theme-based approach offered a distinct advantage in helping students recognize and utilize linguistic cues in reading. Regarding the total reading test score, students in the theme-based model (M = 23.74, SD = 2.77) achieved significantly higher scores than those in the adjunct model (M = 22.20, SD = 2.54; U = 415.50, Z = −2.33, p = .020). The difference between the theme-based and sheltered model (M = 22.06, SD = 3.70) approached statistical significance (U = 452.50, Z = −1.89, p = .059), while the adjunct and sheltered models were not significantly different (U = 608.50, Z = −0.05, p = .962). These findings collectively suggest that the theme-based model had a stronger positive impact on overall English reading proficiency, particularly in equipping students to interpret linguistic features in texts, whereas differences in interpreting implicit information were minimal across instructional models.

Table 3. Mann–Whitney U Test Statistics of English Reading Test

Reading Skill	Statistics	Model 1-Model 2	Model 1-Model 3	Model 2-Model 3
Understanding Implicit Information	Mann-Whitney U	581.500	559.000	590.500
	Wilcoxon W	1211.500	1189.000	1220.500
	Z	-0.370	-0.638	-0.264
	Asymp. Sig. (2-tailed)	0.711	0.523	0.792
Using Linguistic Features to Understand Texts	Mann-Whitney U	315.500	367.000	607.500
	Wilcoxon W	945.500	997.000	1237.500
	Z	-3.634	-3.058	-0.061
	Asymp. Sig. (2-tailed)	0.000	0.002	0.952
Total Reading Test Score	Mann-Whitney U	415.500	452.500	608.500
	Wilcoxon W	1045.500	1082.500	1238.500
	Z	-2.329	-1.890	-0.047
	Asymp. Sig. (2-tailed)	0.020	0.059	0.962

Note: Model 1 = Theme-Based Model; Model 2 = Adjunct Model; Model 3 = Sheltered Model.

In addition to comparisons among the three CBI classes, Wilcoxon Signed Ranks Tests were conducted to examine within-group differences in English reading performance from pretest to posttest for each instructional model. The results are presented in Table 4. Across all three groups, statistically significant improvements were observed in the total reading test scores, indicating that each CBI model positively contributed to overall reading proficiency: theme-based model ($Z = -5.13$, $p < .001$), adjunct model ($Z = -5.18$, $p < .001$), and sheltered model ($Z = -4.94$, $p < .001$). For the sub-skill "Understanding Explicit Information," all three models also showed significant gains: theme-based ($Z = -3.66$, $p < .001$), adjunct ($Z = -4.65$, $p < .001$), and sheltered ($Z = -3.98$, $p < .001$). Similarly, significant progress was detected in "Understanding Implicit Information" within the theme-based ($Z = -3.61$, $p = .004$), adjunct ($Z = -4.81$, $p < .001$), and sheltered models ($Z = -4.49$, $p = .004$), suggesting that students became better able to infer meaning beyond surface-level details. However, the pattern for "Using Linguistic Features to Understand Texts" was more nuanced. Significant improvements were observed in the adjunct ($Z = -4.29$, $p < .001$) and sheltered ($Z = -3.39$, $p < .001$) models, but the theme-based model did not reach statistical significance ($Z = -0.38$, $p = .382$), despite having the highest posttest mean. This may suggest a ceiling effect or possible variations in the instructional focus or student responsiveness across groups. Together, these within-group results affirm the overall effectiveness of CBI in enhancing different dimensions of L2 reading, while also highlighting model-specific strengths and limitations.

Table 4. Wilcoxon Signed Ranks Test Statistics of English Reading Test

Class	Test Statistics[a]	Reading Skill (Pretest-Posttest)			
		Understanding Explicit Information	Understanding Implicit Information	Using Linguistic Features to Understand Texts	Total Reading Test Score
Theme-Based Model	Z	-3.663[b]	-3.605[b]	-4.539[b]	-5.133[b]
	Asymp. Sig. (2-tailed)	0.000	0.004	0.382	0.000
Adjunct Model	Z	-4.648[b]	-4.812[b]	-4.287[b]	-5.184[b]
	Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000
Sheltered Model	Z	-3.975[b]	-4.487[b]	-3.389[b]	-4.942[b]
	Asymp. Sig. (2-tailed)	0.000	0.004	0.382	0.000

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

The quantitative analysis revealed that all three CBI instructional models (theme-based, adjunct, and sheltered) were effective in enhancing students' English reading proficiency, as evidenced by statistically significant gains in total reading scores and most sub-skills from pretest to posttest. At baseline, no significant differences were found among the groups, confirming comparability prior to intervention. Post-intervention, the theme-based model demonstrated a distinct advantage. While all groups improved in interpreting explicit and implicit information, only the theme-based group significantly outperformed the others in the sub-skill using linguistic features to understand texts and achieved the highest overall reading scores. Pairwise comparisons confirmed that the theme-based group's advantage in these areas was statistically significant compared to the adjunct and sheltered models. Within-group comparisons showed that all models led to significant gains in understanding explicit and implicit information and in total scores. Interestingly, although the theme-based group achieved the highest mean in using linguistic features, the improvement was not statistically significant within that group, suggesting a possible ceiling effect or early acquisition advantage. Overall, the findings suggest that while CBI in general improves reading outcomes, the theme-based approach may be particularly effective in promoting deeper linguistic awareness and overall reading performance.

4.2 Qualitative Findings

4.2.1 Inferring Meaning through Guided Interpretation and Content Framing

Across all three instructional models, students' ability to understand implicit information was shaped by the degree to which teachers facilitated inferencing during reading activities. This theme draws on both observational fieldnotes and post-intervention interviews, highlighting differences in how meaning beyond surface-level text was scaffolded in each class. In the theme-based model, for example, classroom observations revealed that teachers frequently paused to ask interpretive questions such as "What do you think the writer is suggesting here?" or "Can you infer the author's intention behind this paragraph?" These prompts encouraged students to move beyond literal comprehension. Interview data confirmed that many students became more attuned to reading between the lines. One participant remarked, *"Our teacher always asked us to think deeply. It was not only about the answer, but why the writer said that."* Another explained, *"Sometimes we were guided to guess the hidden meaning. It helped me understand more."*

In the adjunct model, inferencing occurred less consistently. Observations showed that most reading tasks focused on factual recall or vocabulary clarification, with occasional interpretive questions appearing as extensions rather than central tasks. Correspondingly,

students in this group expressed mixed experiences. While some noted that they were sometimes asked to "guess meanings," others reported limited opportunities to practice this skill. As one student put it, *"We had some hard sentences to explain, but not much about what the whole paragraph was really saying."* In contrast, in the sheltered model, classroom observations showed that instruction largely centered on content understanding, with little attention to implicit meaning. Teachers focused on conveying subject knowledge, often paraphrasing or translating difficult parts of the text rather than inviting student interpretation. This was reflected in the interviews. Students commonly described reading as a tool to "get the information" rather than as an activity requiring critical engagement. A student stated, *"Mostly we just tried to understand the topic. We didn't talk about what is hidden or implied."*

Collectively, this theme helps explain the modest and statistically non-significant differences among the three groups in the quantitative findings for the sub-skill of understanding implicit information. While the theme-based model integrated more opportunities for inferencing, these were not sufficient to produce consistently large between-group effects, especially given the varying levels of student responsiveness and cognitive challenge across contexts.

4.2.2 Differential Attention to Language Form Shapes Linguistic Sensitivity

Students' ability to use linguistic features to understand texts, such as cohesive devices, grammatical patterns, and lexical clues, was strongly mediated by the extent and manner in which teachers directed attention to language form during reading instruction. This theme emerged from triangulating observational data with student interview accounts and offers a qualitative explanation for the significant between-group differences observed in the posttest, particularly the superior performance of the theme-based group.

In the theme-based model, teachers were frequently observed drawing attention to how language constructs meaning. During reading sessions, instructors explicitly unpacked sentence structure, highlighted cohesive markers (e.g., *however*, *in contrast*), and explained the functions of relative clauses, noun phrases, and logical connectors. This practice was often framed around the reading task, linking form to function. A typical example from fieldnotes recorded a teacher asking, *"What does 'which' refer to in this sentence? Does it help you understand what the author is arguing?"* In interviews, students expressed that such moments helped them understand how grammar and vocabulary contributed to meaning. One student shared, *"The teacher told us to pay attention to words like 'although' or sentence order. I didn't notice these before, but now I can use them to guess meaning."*

In contrast, the adjunct model showed sporadic attention to linguistic features. While some language elements were discussed, particularly in pre-reading vocabulary stages, form-focused instruction was typically decontextualized and not systematically connected to reading comprehension. Observations noted that grammar points were sometimes addressed in isolation, often during brief digressions. Interviews echoed this inconsistency. One student remarked, *"We had language focus sometimes, but it was not really linked to the reading part. It was like a different lesson."* Another noted, *"I remember some grammar, but I'm not sure how to use it for reading."*

In the sheltered model, focus on linguistic features was minimal. Lessons emphasized content knowledge and topical understanding, with teachers simplifying text passages through paraphrasing or explanations but rarely commenting on how language conveyed meaning. Observations indicated that students were seldom asked to analyze sentence structure or lexical cohesion. Interviews confirmed this pattern. One sheltered group student commented, *"We mostly just talked about the topic, like what the reading was about. The language part was skipped."*

This theme provides a clear explanatory lens for the quantitative findings, where the theme-based model significantly outperformed the other two groups on the sub-skill *Using Linguistic Features to Understand Texts*. The rich and consistent integration of form-focused instruction in the theme-based class appeared to sharpen students' linguistic awareness, while the more fragmented or absent attention in the other models limited students' development in this area.

4.2.3 Depth of Discourse Mediation Facilitates Inference-Making

Students' proficiency in understanding implicit information, such as identifying implied meanings, author stance, or underlying assumptions, was closely linked to how teachers mediated discourse during reading instruction. This theme captures how varying depths of teacher-led interpretation shaped students' inferential reasoning across the three CBI models and helps explain why statistically significant differences were observed at the group level, yet pairwise comparisons failed to detect clear distinctions.

In the theme-based model, classroom observations revealed that teachers frequently engaged students in higher-order questioning and reflective discussion. During text analysis, instructors posed interpretive questions such as *"Why do you think the author chose this example?"* or *"What is not directly said but implied here?"* Students were encouraged to go beyond literal comprehension and explore subtext and author intention. Interviewees confirmed the value of this approach. One student explained, *"The teacher asked us to guess what the writer was trying to say, even if it's not written. At first I didn't know how, but with practice I improved."*

In the adjunct model, mediation of implicit meaning was somewhat present but less consistent. Teachers occasionally introduced inferential questions but often followed a fixed list of comprehension questions without elaboration. Observational data noted that when inferencing was required, it was usually scaffolded through vocabulary glosses or leading prompts. A student from this class shared, *"Sometimes we had to think more deeply, but it depended on the teacher or the activity. It wasn't every time."* Another added, *"I could understand the basic meaning, but not always what the writer wanted us to feel or assume."*

In the sheltered model, discourse mediation was heavily content-oriented, with limited emphasis on reading between the lines. Teachers focused on explaining factual information or key ideas explicitly stated in the text. Observations recorded few instances where students

were invited to explore implied meanings. Interviews echoed this pattern. One student noted, *"Most of the time we just talked about what the text says directly. I didn't learn how to guess the hidden meaning."*

These findings help interpret the quantitative results: while the Kruskal–Wallis test identified significant group differences in posttest scores for *understanding implicit information*, no single pairwise comparison reached significance, possibly due to overlapping instructional features. The theme-based model showed more consistent mediation of inference-making, but both the adjunct and sheltered models occasionally engaged with this skill, albeit inconsistently. This blurred boundary across models is reflected in the marginal quantitative distinctions.

4.2.4 Explicit Attention to Linguistic Cues Enhances Textual Awareness

This theme highlights the extent to which teachers across the three CBI models drew students' attention to grammatical structures, discourse markers, and cohesive devices as tools for comprehension. These linguistic features, though often overlooked in content-heavy instruction, proved essential for students' ability to interpret sentence relationships, resolve referents, and track text logic. The presence or absence of this linguistic focus directly aligns with the quantitative findings, where the theme-based group significantly outperformed the others on this sub-skill.

In the theme-based model, classroom observations recorded frequent and deliberate emphasis on linguistic features during reading tasks. Teachers regularly paused to highlight connectors (e.g., *however, as a result*), cohesive devices (e.g., pronoun referencing), and clause relationships (e.g., cause-effect, contrast). Students were also asked to paraphrase or restate passages using those features, reinforcing their function in textual coherence. One student reflected, *"My teacher always explained the words that join ideas or show cause and effect. That helped me understand how parts of the reading connect."* Another noted, *"We even did exercises to change sentence structures. It helped me follow the argument."*

In the adjunct model, while some attention was given to language form, it was often incidental and secondary to the content. Observations revealed that instructors occasionally mentioned useful phrases or structures, usually when students encountered confusion, but such attention was reactive rather than pre-planned. One student mentioned, *"Sometimes we looked at phrases that help the meaning, but it was not the focus. It happened more when someone asked."* Another said, *"I remember learning words like 'moreover' or 'thus,' but only because they were in the glossary."*

In contrast, the sheltered model placed minimal emphasis on linguistic features as tools for comprehension. Instruction centered predominantly on factual content delivery and conceptual clarification, with little reference to how language shaped meaning. Observers noted that while students read academic texts, there was no sustained discussion of how the grammar or discourse structure affected interpretation. A student in this group stated, *"We read the article and talked about the ideas, but the teacher didn't teach us the sentence links or the grammar patterns."*

These classroom practices map clearly onto the quantitative results: the theme-based group scored significantly higher in the use of linguistic features to understand texts compared to both the adjunct and sheltered groups. The qualitative data affirm that this advantage stems not merely from exposure to content but from the integration of language instruction into content teaching, which is a core principle of effective CLIL implementation.

5. Discussion

The findings of this mixed-methods study revealed distinct patterns across the three CBI instructional models in enhancing English reading proficiency among Chinese EFL university students. While all three models led to significant gains from pretest to posttest, differences emerged in the extent and nature of improvement, particularly in the sub-skills of understanding implied information and using linguistic features to understand texts. These patterns were further illuminated by qualitative data from classroom observations and interviews, which offered insights into instructional design, teacher practices, and student experiences across the three models.

Quantitative analysis showed that students in the theme-based model significantly outperformed their counterparts in the other two models in using linguistic features to understand texts and achieved higher overall reading proficiency. This finding aligns with prior research suggesting that when language learning is integrated meaningfully with content (Hu et al., 2025; Zhang et al., 2024), students are more likely to develop metalinguistic awareness and lexical sensitivity, which are essential for advanced reading comprehension (Amin et al., 2021; Fikni et al., 2024; Khusniyah & Wadi, 2020). The qualitative data reinforced this interpretation: students in the theme-based class reported more frequent engagement with morpho-syntactic cues and textual connectors, while observations revealed that their teacher explicitly modeled how linguistic markers signal meaning. In contrast, in the adjunct and sheltered models, linguistic features were either addressed separately from content (adjunct) or oversimplified through teacher paraphrasing (sheltered), which may have limited students' exposure to authentic, challenging input. This aligns with findings from Hu and AlSaqqaf (2024) and Suvonova (2023), who highlighted the risk of linguistic oversimplification in content-dominant instruction. As a contribution, this study isolates model effects, links observed classroom practices to subskill-specific outcomes, and refines CBI theorizing by showing that gains in reading proficiency are especially sensitive to integrated, theme-based design.

Unexpectedly, although all three groups improved significantly in understanding implicit information, no statistically significant differences were found among them in the posttest. This contradicts earlier claims that tightly integrated models such as theme-based CBI are more effective in developing inferencing skills (Sariani et al., 2022; Snow & Brinton, 2023). Grabe and Yamashita (2022) and Xiong

and Jiang (2020) attribute the advantage of integrated models to consolidated topical schema, repeated discourse patterns, and guided strategy instruction that should specifically boost bridging and elaborative inference. Under our controlled materials and equal time-on-task, however, these putative advantages did not yield between-group differences, suggesting that inferencing may be constrained more by inferential load and prior knowledge than by model structure per se. Our findings thus refine the literature by indicating that integration is sufficient for within-group gains but not a unique driver of between-group effects without targeted inference-strategy dosing and texts with higher inferential density. The qualitative findings help explain this outcome. Observations confirmed that inferencing tasks were inconsistently implemented across all three models (Masharipova, 2022): in the theme-based and sheltered classes, teachers often provided scaffolds that led students directly to answers, while in the adjunct model, inferencing was more content-driven than language-driven. Additionally, interview data suggest that while students developed awareness of implicit meaning, they struggled to articulate how they arrived at such understandings, indicating that gains may have been more intuitive than systematic. This complexity resonates with Nation's (2009) distinction between declarative and procedural knowledge in L2 reading.

Quantitative findings also indicated significant within-group gains for all models across all three sub-skills, but the theme-based group did not show a statistically significant gain in using linguistic features, despite having the highest posttest mean. This apparent contradiction may be due to a ceiling effect, whereby students in the theme-based model had already developed high proficiency during the pretest or early stages of the intervention, limiting room for measurable growth (S. Li, 2020). Interview data lend partial support to this interpretation: some students reported that they had become accustomed to textual markers due to earlier exposure to integrated tasks, suggesting that the instructional model may have reinforced rather than newly developed these skills (Chekol et al., 2023). This nuanced outcome underscores the importance of considering both mean differences and effect trajectories in interpreting intervention impact (Hu & AlSaqqaf, 2024). It also echoes Zhang and Tang's (2024) call for closer attention to learners' starting proficiency levels in vocabulary and discourse sensitivity when evaluating reading outcomes in CBI.

Another notable distinction emerged in the patterns of classroom interaction and student autonomy. The theme-based class was characterized by frequent peer interaction, task-based negotiation, and teacher scaffolding that encouraged student reflection on both content and language. This was consistent with the model of form-focused instruction within communicative contexts (Brinton & Snow, 2017; Snow & Brinton, 2023), which posits that drawing learners' attention to form during meaning-focused tasks enhances long-term retention and transferability. Conversely, the adjunct model saw limited cross-disciplinary coordination (Karimi & Ghorbanchian, 2022), with students often needing to reconcile content and language learning independently. While some autonomy was fostered, interviews indicated that many students felt uncertain about expectations and struggled to transfer vocabulary knowledge into academic reading tasks. In the sheltered model, although students appreciated the simplified input and guided practice, their inferencing and linguistic reasoning were often teacher-dependent (Echevarria & Graves, 2014; Latif, 2024), potentially hindering the development of strategic reading abilities. These patterns support the findings of Cheng (2024) and Zhang et al. (2024), who emphasized the role of instructional alignment in facilitating deeper cognitive engagement in CBI settings. However, the findings contradict prior reports that theme-based or sheltered models uniformly outperform alternatives on higher-order reading (e.g., S. Li, 2020; Sariani et al., 2022). Here, advantages emerged only when instruction paired integration with explicit form–function modeling and learner-driven inference, not from model type alone.

The observed variation in outcomes across the three CBI models also raises important questions about the role of instructional coherence and the distribution of cognitive demands. In the theme-based model, the simultaneous activation of content and language processing may have increased learners' cognitive engagement (Fikni et al., 2024; Zhang & Tang, 2024) in a way that aligns with the Cognitive Load Theory (Sweller, 1988). By embedding language instruction directly within meaningful content contexts, the model may have facilitated germane cognitive load (Suvonova, 2023), which supports schema construction rather than overloading working memory with disconnected information. In contrast, the adjunct model's separation of language and content instruction appeared to shift the cognitive burden onto students (Chekol et al., 2023), who were required to independently reconcile terminology and conceptual structures across two courses. Similarly, the sheltered model's simplified linguistic input and teacher-led questioning may have reduced intrinsic cognitive load but did so at the expense of student agency and critical inferencing opportunities (Latif, 2024). These distinctions suggest that it is not merely the presence of content or language instruction that matters, but the pedagogical design of their integration, a principle echoed in the broader literature on bilingual and integrated instruction (Hu & Zhang, 2024).

6. Conclusion

This study examined the differential effects of three CBI models (theme-based, adjunct, and sheltered) on Chinese EFL university students' English reading proficiency, focusing on the sub-skills of understanding explicit information, understanding implicit information, and using linguistic features to understand texts. Quantitative findings revealed that while all models produced significant improvements from pretest to posttest, the theme-based model led to notably higher gains in using linguistic features and overall reading performance. Interestingly, no significant differences were found among the groups in their ability to understand implicit information. Qualitative data from classroom observations and interviews provided further insight, highlighting key differences in instructional coherence, teacher practices, and the nature of language–content integration that help explain these varied outcomes. As a contribution to knowledge, this study isolates model effects while holding textbook, contact hours, and assessment constant and disaggregates reading into explicit, implicit, and linguistic-feature sub-skills, showing that theme-based design is particularly conducive to strengthening cohesion- and syntax-supported comprehension. It also suggests that gains in implicit-information processing may be less sensitive to model type than to task design and inferential load, refining where CBI implementations should target change.

The findings carry important implications for curriculum designers, teacher educators, and classroom practitioners. First, the demonstrated effectiveness of the theme-based model suggests that tightly integrated content-and-language instruction, where language is not treated as a separate component, can more effectively support students' lexical awareness and discourse-level reading skills. Second, the observed instructional misalignments in the adjunct and sheltered models point to the need for coordinated planning and pedagogical training to avoid fragmented or oversimplified instruction. Teacher development programs should emphasize not only content expertise or language pedagogy in isolation but also how these can be integrated meaningfully. Finally, the role of interactional scaffolding and student agency observed in the theme-based class underscores the pedagogical value of promoting dialogic learning and form-focused attention within meaning-rich tasks.

Despite its contributions, this study is not without limitations. First, the sample was limited to three intact classes from a single university in China, which may constrain the generalizability of the findings to other institutional, regional, or national contexts. Future studies should therefore include more diverse samples from multiple institutions to allow for broader comparisons and improve external validity. Second, although the study employed a robust mixed-methods design, it relied primarily on classroom observations and post-intervention interviews to capture qualitative insights. These methods, while valuable, may not have fully captured the complexity of in-the-moment cognitive and interactional processes during reading. Incorporating real-time data collection methods such as stimulated recall interviews or lesson transcripts across different stages of instruction could provide a more comprehensive picture. Lastly, the study focused only on immediate posttest effects, leaving the long-term sustainability of the observed reading gains unknown. Future research should include delayed posttests to assess retention and the transferability of learned reading skills over time. Addressing these limitations would not only strengthen the empirical base for evaluating different CBI models but also offer more actionable insights into refining content-integrated pedagogies for L2 reading instruction.

Acknowledgments

We would like to thank the research participants who took the time to participate in this study.

Authors' contributions

Lu Zhang contributed to the conceptualization of the study, data collection, and drafting of the manuscript. Hanita Hanim Ismail provided supervision, methodological guidance, and critical revisions to strengthen the theoretical framing and analysis. Nur Ainil Sulaiman contributed to data analysis, interpretation of findings, and refinement of the manuscript for clarity and coherence. All authors reviewed and approved the final version of the manuscript.

Funding

No external funding was received by the authors.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Amin, S., Salija, K., & Basri, M. (2021). Students' perception toward the integrated content-based reading instructional material. *Asian Journal of Humanities and Social Studies*, *9*(4), 138–150. https://doi.org/10.24203/ajhss.v9i4.6690

Birch, B. M. (2014). *English L2 Reading: Getting to the Bottom* (2nd ed.). Routledge.

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage.

Brinton, D., & Snow, M. A. (Eds.). (2017). *The Content-based Classroom: New Perspectives on Integrating Language and Content*. University of Michigan Press.

Chandler, P. T. (2020). *Implementing The SIOP Model to Support English Language Learners* [Doctoral thesis, University of New England]. https://dune.une.edu/theses/315

Chekol, A., Shiferie, K., & Teshome, S. (2023). Effects of using adjunct model of content-based instruction on students' technical report writing performance. *Cogent Education*, *10*(1), Article 2214427. https://doi.org/10.1080/2331186X.2023.2214427

Cheng, L. (2024). Chinese Storytelling Theme-oriented in English CBI Teaching and Research. In *Proceedings of the 2024 3rd International Conference on Comprehensive Art and Cultural Communication* (pp. 77-82). Atlantis Press. https://doi.org/10.2991/978-2-38476-287-3_10

College Foreign Language Teaching Steering Committee of the Ministry of Education. (2020). *大学英语教学指南 [Guidelines for College English Teaching]*. Higher Education Press.

Echevarria, J., & Graves, A. (2014). *Sheltered content instruction: Teaching English learners with diverse abilities*. Pearson Education.

Fikni, Z., Wathoni, H., Nurhaliyani, N., & Husnu, M. (2024). Assessing Content Based Instruction's Impact on English Reading Skill. *Jurnal Pendidikan Bahasa Inggris Indonesia*, *12*(2), 129–137. https://doi.org/10.23887/jpbi.v12i2.3841

Grabe, W., & Yamashita, J. (2022). *Reading in a Second Language: Moving from Theory to Practice*. Cambridge University Press.

Headley, M. G., & Clark, V. L. P. (2019). Multilevel mixed methods research designs: Advancing a refined definition. *Journal of Mixed Methods Research*, *14*(2), 145–163. https://doi.org/10.1177/1558689819844417

Hennink, M. M., & Kaiser, B. N. (2020). *Saturation in qualitative research*. Sage.

Hu, H., Hashim, H., & Li, X. (2025). Bridging language and learning: A CLIL-ized TMI model for international students in traditional Chinese medicine. *Journal of Medical Education Development*, *18*(1), 140-146. https://doi.org/10.61186/edcj.18.1.140

Hu, H., Said, N. E. M., & Hashim, H. (2022). Killing two birds with one stone? A study on achievement levels and affective factors in content and language integrated learning (CLIL). *International Journal of Learning, Teaching and Educational Research*, *21*(4), 150-167. https://doi.org/10.26803/ijlter.21.4.9

Hu, H., & Zhang, L. (2024). Towards a connected CLIL research agenda: Bridging gaps and charting new horizons. In G. Zeybek (Ed.), *Policy Development, Curriculum Design, and Administration of Language Education* (pp. 216–236). IGI Global. https://doi.org/10.4018/979-8-3693-3645-8.ch009

Hu, K., & AlSaqqaf, A. (2024). Conceptualizing the Integration of a 3As Approach Content-Based Instruction into Teaching Business English Speaking to Chinese EFL Undergraduates. *International Journal of Academic Research in Business and Social Sciences*, *14*(2), 367-382. https://doi.org/10.6007/IJARBSS/v14-i2/20728

Huang, J., Lin, R., Wang, Q., Su, M., & Zhao, B. (2022). *大学英语四级阅读专项突破 [Special breakthrough in College English Test Band 4 reading]*. Tsinghua University Press.

Karimi, M., & Ghorbanchian, E. (2022). Effects of Adjunct Model of Instruction on EAP Learners' Reading Comprehension Skill. *Theory and Practice of Second Language Acquisition*, *8*(2), 1-20. https://doi.org/10.31261/TAPSLA.10432

Khusniyah, N. L., & Wadi, A. S. (2020). Investigating Content-Based Reading Instruction In Promoting Students' Reading Comprehension: A Classroom Action Research. *Research and Innovation in Language Learning*, *3*(2), 96-106. https://doi.org/10.33603/rill.v3i2.3206

Komarawan, Y. (2024). Implementing reading to learn pedagogy in CLIL lessons: The case of an Indonesian secondary school. *SEAQIL Journal of Language Education*, *3*(1), 125–136. https://doi.org/10.70046/sjle.3.1.125--136

Latif, O. L. I. A. (2024). Using content-based sheltered instruction observation protocol (SIOP) for developing faculty of education second year English majors' language proficiency. *Journal of Research in Education and Psychology*, *39*(2), 265-319. https://doi.org/10.21608/MATHJ.2024.286741.1440

Li, S. (2020). The Impact of CBI on English Majors' Vocabulary Learning. In C. J. Hyung, Z. Jin, L. F. Ying, & C. F. Peng (Eds.), *Proceedings of the 2020 International Conference on Language, Art and Cultural Exchange* (pp. 102-105). Atlantis Press.

Li, X., & Fu, H. (2021). Needs Analysis on ESP Course for Business and Engineering Students in a Chinese Local University. *Creative Education*, *12*, 1387-1396. https://doi.org/10.4236/ce.2021.126105

Li, Y. (2020). *Educational Change Amongst English Language College Teachers in China: Transitioning from Teaching for General to*

Academic Purposes. Springer.

Lyster, R., & Ballinger, S. (2011). Content-based language teaching: Convergent concerns across divergent contexts. *Language Teaching Research*, *15*(3), 279-288. https://doi.org/10.1177/1362168811401150

Masharipova, F. J. (2022). Techniques for enhancing reading competence on the CBI/STEM domain. *Foreign Languages in Uzbekistan*, *4*(45), 98-109. https://doi.org/10.36078/1666329625

Ministry of Education of the People's Republic of China, & National Language Commission of the People's Republic of China. (2018). *China's Standards of English Language Ability*. https://cse.neea.edu.cn/res/ceedu/1811/6bdc26c323d188948fca8048833f151a.pdf

Mohaideen, M. S. H., Ismail, H. H., & Rashid, R. A. (2020). The use of local literary texts as reading materials in English language classrooms: An analysis of teachers' perspectives. *International Journal of Learning, Teaching and Educational Research*, *19*(11), 127-144. https://doi.org/10.26803/ijlter.19.11.8

Nation, I. S. P. (2009). *Teaching ESL/EFL Reading and Writing*. Routledge.

Ponna, W., Buehler, R., & Zhou, G. (2024). Exploring the Effectiveness of Concept-Based Instruction Versus Deductive Methods in Teaching English Grammar. *Advances Educational Innovation*, *1*(2), 43-51. https://doi.org/10.69725/aei.v1i2.146

Sadeghi, K. (2021). *Assessing Second Language Reading: Insights from Cloze Tests*. Springer Cham.

Sariani, S., Yaningsih, Y., & Rozi, F. (2022). Implementing content-based instruction (CBI) on EFL student's writing through technology-enhanced language teaching. *The Asian Institute of ResearchEducation Quarterly Reviews*, *5*(1), 174-184. https://doi.org/10.31014/aior.1993.05.01.429

Snow, A. M., & Brinton, D. M. (2023). *Content-based instruction: What every ESL teacher needs to know*. University of Michigan Press.

Suvonova, D. K. (2023). General overview of content-based instruction approach (CBI) in teaching foreign languages. *Golden Brain*, *1*(17), 374-380. https://doi.org/https://researchedu.org/index.php/goldenbrain/article/view/4102

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Xiong, J., & Jiang, X. (2020). Reflections on foreign language education policy reform in China under the background of "One Belt and One Road" initiative *Sino-US English Teaching*, *17*(5), 149–153. https://doi.org/10.17265/1539-8072/2020.05.002

Yu, Y., Deng, N., Xu, M., & Liang, S. (2024). From Classroom to Marketplace: Empowering Business English Students with Cross-Border Product Digital Storytelling. *Journal of Research in Vocational Education*, *6*(8), 73-79. https://doi.org/10.53469/jrve.2024.06(08).16

Zhang, L., Ismail, H. H., & Sulaiman, N. A. (2024). Content-based instruction in language education: A bibliometric review. *Forum for Linguistic Studies*, *6*(4), 388–403. https://doi.org/10.30564/fls.v6i4.6845

Zhang, M., & Tang, L. (2024). A Review of Content-Based Instruction on Critical Thinking in College English Reading Instruction. *International Journal of Education and Humanities*, *16*(1), 177-181. https://doi.org/10.54097/3xwwds97

Zheng, L., Ismail, H. H., Hashim, H., & Ali, Z. B. (2025). Storytelling in English Language Education in China: A Systematic Review of Empirical Research from the Past Decade (2014–2024). *Forum for Linguistic Studies*, *7*(2), 280-295. https://doi.org/10.30564/fls.v7i2.8314