Jurnal Pendidikan Edutama
Volumes 13 Number 2 July 2026
P-ISSN: 2339-2258 | E-ISSN: 2548-821X
IKIP PGRI Bojonegoro

EFL Students' Difficulties in Applying Lexical Inferencing Strategies in Reading Authentic Literary Texts

Muhammad Syafrizal Amrul[1]* Intan Permata Hapsari[2]
*[1.2]Universitas Negeri Semarang

*[1]muhsyafrizal38@students.unnes.ac.id; [2]intan74@mail.unnes.ac.id

***Corresponding Author**

Keywords

Lexical Inferencing, Authentic Literary Texts, EFL Reading, Nassaji Taxonomy, Qualitative Research

Abstract

While lexical inferencing helps EFL learners deduce words, it frequently fails with authentic literature due to figurative language and cultural nuances. This qualitative study investigated the inferencing difficulties of English Education students reading familiar and unfamiliar literary texts. Data from five purposively selected students were gathered via task-based questionnaires and interviews, then analyzed using Miles et al.'s (2018) interactive model and Nassaji's (2003) taxonomy. Results showed that despite using various strategies, students consistently encountered failures. They often fell for false lexical anchors caused by polysemy and rarely self-corrected. Furthermore, strong storyline engagement led to "narrative prediction," prompting students to guess meanings based entirely on plot logic before attempting linguistic analysis. Because these errors persisted across both texts, this study concludes that inferencing breakdowns stem from literary prose complexity rather than text unfamiliarity.

Introduction

When an EFL student encounters the word *tumultuously* in a sentence that also contains the word *rose*, which word takes priority? The answer, this study finds, is almost never the word they are supposed to be reading. The observation is small, but the problem behind it is not. EFL learners rely on lexical inferencing (deriving word meaning from context, linguistic cues, and

prior knowledge) as a core reading strategy (Grabe & Stoller, 2019; Nation, 2022). But in literary texts, this strategy breaks down in ways that standard reading research has not yet fully mapped. Processing unknown words is already cognitively expensive: it increases load, exhausts working memory, and causes readers to lose the thread (Elgort et al., 2018). In literary texts, the problem runs deeper. Even familiar words can lead readers astray. Managing unknown words requires the use of an independent strategy (Ali, 2020). However, in literary reading, that is only half the problem. The other half is recognizing when the text itself is working against the strategy.

Lexical inferencing sits at the center of this problem. Nassaji (2003) defines it as deriving word meaning from contextual information, linguistic cues, and prior knowledge. He organizes it into five strategy categories: contextual and semantic cues, morphological analysis, syntactic cues, world knowledge, and interlingual transfer. What makes this taxonomy useful for the present study is not just that it names five strategies, but that each category points to a different point in the inferencing process where failure can occur. Nassaji's framework makes it possible to ask not only whether students failed, but exactly where and why. In L1 reading research, morphological analysis has also been shown to predict reading comprehension in young learners (Levesque et al., 2017), though the mechanisms differ from those in L2 contexts. Morphological analysis in particular (breaking unfamiliar words into their constituent parts to recover meaning) has been directly linked to lexical inferencing success and vocabulary development in L2 learners (Zhang & Koda, 2012). When the process works, it does more than enable immediate comprehension: it builds vocabulary over repeated encounters (Hassanzadeh et al., 2019).

None of this works as smoothly in authentic literary texts. While previous studies in the Indonesian context have shown that utilizing narrative texts based on familiar local culture can significantly aid EFL learners in grasping storyline components and improving comprehension (Fauzi et al., 2024), authentic texts, including literary prose, are written for native readers rather than language learners. This means they completely lack any built-in comprehension scaffolding (Nation, 2022). In English Education programs, they appear in courses like Literary Reading as required material, which means students cannot simply avoid them. Literature has genuine value for language development: it exposes learners to authentic target-language use and pushes them toward critical and intercultural thinking (Bobkina & Stefanova, 2016). However, it does this through figurative language, deliberate polysemy, and cultural allusion. These are features that resist any single stable reading of a word (Alkhammash, 2022; Carrol et al., 2018). Academic texts are usually predictable in structure; literary texts are not, and that difference matters for inferencing. Bunparit and Chinokul (2018) found that EFL learners already struggled with inferencing in academic texts, relying on a narrow range of strategies and frequently arriving at wrong meanings. In literary texts, where the contextual clues surrounding an unknown word are themselves figurative or culturally specific, that struggle intensifies.

Yet the existing research has not seriously engaged with the literary context. Yang's et al. (2023) bibliometric analysis of 472 lexical inferencing articles from the Web of Science Core Collection found that nearly all of them situate inferencing in general or academic reading.

Literary texts barely appear. The studies that do exist tend toward quantitative designs: correlating inferencing ability with vocabulary scores, or testing instructional interventions on graded texts (Rutamornchai & Tepsuriwong, 2022). What is missing is qualitative inquiry. The field requires studies that explore what actually happens when a learner sits with a literary sentence and tries to work out an unfamiliar word. What goes wrong, and why? The field has largely treated reading texts as essentially interchangeable. Yang et al. (2023) document this exact tendency, noting that the vast majority of lexical inferencing research makes no distinction between literary and academic reading contexts, despite the fundamentally different cognitive demands these text types impose on the learner.

The pedagogical stakes are specific. English Education students are prospective teachers who will one day guide learners through the same literary texts they are struggling with now. Knowing exactly how and why their inferencing breaks down when facing literary polysemy or cultural allusions is essential. It equips these future educators with the precise knowledge needed to design effective pedagogical interventions. Without it, instructors are left with generic advice: read more, expose students to more authentic texts (Grabe & Stoller, 2019). That advice is not wrong, but it is not enough. Therefore, the novelty of this research lies in its qualitative exploration of lexical inferencing specifically within authentic literary texts, a domain heavily underrepresented in current literature. Therefore, the novelty of this research is twofold. First, it addresses a critical gap by qualitatively investigating lexical inferencing failure within authentic literary short stories, a context heavily underrepresented in the literature (Yang et al., 2023). Second, moving beyond quantitative outcomes, this study maps the precise cognitive pathways and breakdown points within Nassaji's (2003) five strategy categories in literary prose. This approach offers a level of process-level granularity previously absent from the field, paving the way for a deeper theoretical understanding of lexical inferencing beyond standard academic reading contexts. Thus, this study asks two questions. First, what specific difficulties do fourth-semester EFL students face when applying lexical inferencing strategies in authentic literary texts? Second, what linguistic and non-linguistic factors contribute to those difficulties?

Method

Research design, Sample, and Procedure

The core questions of this study regarding what goes wrong and why, when specific students try to infer specific words, cannot be adequately answered through quantitative metrics. While experimental studies have provided valuable insights into how word frequency and recurrence in audio-visual input affect incidental vocabulary acquisition (Peters & Webb, 2018) and the visual attention and processing behaviors exhibited when encountering unknown words during reading (Pellicer-Sánchez, 2016), understanding the actual reasoning process, specifically what students noticed, what they assumed, and where they got stuck, requires direct access to their thinking. A qualitative descriptive approach is appropriate in this study

because the goal is to describe a phenomenon in depth within a specific context, not to generalize across populations (Rose et al., 2019).

Participants were fourth-semester undergraduate students enrolled in the Literary Reading course within the English Education study program at a university in Semarang City. The two task-based questionnaires were administered to the full class of 26 students. From this group, five were selected for the interview phase. They were chosen not randomly, but purposively (Campbell et al., 2020). The aim was to find participants whose written responses showed clear, varied, and traceable patterns of inferencing difficulty, and whose explanations were detailed enough to pursue further. One participant with a perfect score was deliberately included as a negative case (Creswell & Poth, 2018) to test whether difficulty arises in the process, even when it does not appear in the answer. These five participants are referred to as P1 through P5 throughout this study. Their individual scores on the questionnaires are reported in the Findings section. All participants signed an informed consent form before the study began and were told they could withdraw at any time. Their names were replaced with pseudonyms throughout analysis and reporting.

Three instruments were used. The first (Test 1) was a task-based questionnaire delivered via Google Forms, built around excerpts from Kate Chopin's The Story of an Hour (a text the class had already read). Five excerpts were drawn from the story, each containing one underlined target word. The five words were selected to represent each of Nassaji's (2003) strategy categories: *tumultuously* (morphological), *afflicted* (contextual/semantic), *forestall* (syntactic), *elusive* (world knowledge), and *intelligence* (interlingual transfer). For each word, students chose from four options and wrote a brief explanation of their reasoning process. Because students already knew the story, Test 1 serves as the baseline: what inferencing looks like when narrative schema is available.

The second (Test 2) used the same format as Test 1, but with different text: O. Henry's The Ransom of Red Chief, which students had not previously encountered. Its five target words were *surreptitiously* (morphological), *somnolent* (contextual/semantic), *acceded* (syntactic), *diatribe* (world knowledge), and *reconnoitre* (interlingual transfer). Pairing the two tests was a deliberate design choice. The goal was not to compare performance across texts but to check whether the same difficulties appeared in both. If a student struggled with the same type of inferencing in a story they already knew and in one they had never read, that struggle could not be blamed on unfamiliarity with the plot. It would have to be attributed to something in the literary language itself. Test 1 was always given first to prevent carryover. During both tests, students were not permitted to use dictionaries, translation tools, or any AI assistance. Because both questionnaires were administered via Google Forms without direct supervision, the possibility that individual participants consulted external resources cannot be entirely ruled out. This represents a limitation inherent to online, unproctored data collection. Both instruments were developed in close alignment with Nassaji's (2003) taxonomy and revised iteratively throughout the supervisory process.

The third instrument was a semi-structured interview conducted individually via WhatsApp text. Rather than following a single shared protocol, each interview was built

specifically around that participant's questionnaire responses: asking about the words they had struggled with, the reasoning they had written down, and the moments where their explanations were vague or puzzling. The questions were meant to go deeper into what had already been said, not to open new topics. Participants could respond in English, Bahasa Indonesia, or both; most used a mix. Responses were translated into English by the researcher for reporting, with the original wording kept in parentheses throughout. WhatsApp text allowed students to respond at their own pace and think before answering. This is a critical advantage when asking people to reconstruct their own cognitive process, which does not come easily in real-time conversation (Kaufmann & Peil, 2020).

Data analysis

Analysis followed the interactive model of Miles et al. (2018), moving between condensing raw data, displaying it in organized form, and drawing conclusions in a process that looped back on itself rather than progressing in a straight line. Coding was deductive: Nassaji's (2003) five strategy categories served as the primary framework, and each piece of data was coded against those categories to identify where in the inferencing process things went wrong. Interview responses were read alongside the questionnaire answers they came from, not as separate data, but as elaboration of what the written responses had left unexplained. Findings were organized into thematic matrices to allow comparison across participants, and patterns from Test 1 and Test 2 were reviewed together to check whether the same difficulties appeared in both textual conditions. Trustworthiness was addressed through triangulation across three instruments, member checking (each of the five participants was asked to review a written summary of the researcher's interpretations of their responses, and none raised substantive disagreement with the characterizations presented), thick description of the research context, and an audit trail documenting analytic decisions (Korstjens & Moser, 2018). Because the study was conducted within the researcher's own program and institution, reflexive journaling was kept throughout the analysis period: a record of assumptions questioned, interpretations reconsidered, and decisions made.

Results and Discussion

Results

The five interview participants answered as follows: P1 answered 4 out of 10 target words correctly, P2 answered 6 out of 10, P3 answered 7 out of 10, P4 answered 8 out of 10, and P5 answered all 10 correctly. These scores reflect the range of inferencing difficulty across the group. Table 1 shows accuracy rates by word and strategy category. The numbers capture what students got right and wrong, and the open-ended responses and interviews reveal the reasoning behind those answers.

Table 1. Accuracy Rates for All Target Words Across Test 1 (T1) and Test 2 (T2)

Word (Test)	Nassaji Strategy	Correct (n=26)	Accuracy (%)
tumultuously (T1)	Morphological	19	73.1%
surreptitiously (T2)	Morphological	24	92.3%
afflicted (T1)	Contextual/Semantic	23	88.5%
somnolent (T2)	Contextual/Semantic	25	96.2%
forestall (T1)	Syntactic	25	96.2%
acceded (T2)	Syntactic	23	88.5%
elusive (T1)	World Knowledge	25	96.2%
diatribe (T2)	World Knowledge	24	92.3%
intelligence (T1)	Interlingual Transfer	18	69.2%
reconnoitre (T2)	Interlingual Transfer	25	96.2%

Morphological Strategy Outcomes

Of the ten target words, *tumultuously* produced the second lowest accuracy in Test 1 (73.1%), with seven of twenty-six students answering incorrectly. The distribution of wrong answers clustered around option B, which is "Peacefully and very quietly". *Surreptitiously* in Test 2 produced higher accuracy (92.3%), though two participants still erred. Questionnaire responses showed that for *tumultuously*, students focused on a single concrete word earlier in the sentence, "rose". The sentence reads: "Her bosom rose and fell tumultuously". These incorrect participants consistently explained that the word 'rose' prompted an immediate association with flowers, softness, or peace. P1 stated: "I thought the word 'rose' here means a flower, and flowers are usually associated with something soft, so I ended up thinking *tumultuously* meant something similar" (original: "Aku berpikir kalo kata 'rose' disini itu artinya bunga mawar... jadinya mikir kata *tumultuously* itu mirip-mirip sama sesuatu yang soft"). P2 similarly admitted: "I was paying attention to 'rose', not to bosom or fell" (original: "Merhatiin 'rose' nya kak, kurang merhatiin yang bosom atau fell"). P3 added an additional layer, noting: "I often see *tumultuously* in poems, so I guessed based on the flower vibe" (original: "Saya kayak sering liat *tumultuously* di puisi-puisi gitu jadi nebak artinya dari identik dan vibe bunga"). In all three cases, participants read 'rose' as a standalone noun rather than as part of the phrasal unit 'rose and fell'. When asked directly whether they had read the phrase as a unit, P2 replied that they had not considered 'fell' at all. Only P5, who answered correctly, articulated a different reading unit: "I wasn't misled because there was 'and fell' right after it" (original: "Nggak kepancing karena ada kata 'and fell' nya sih kak"). P5 also noted that the stacked suffixes of *tumultuously*, particularly the -uously ending, created an initial visual hesitation. P1's interview revealed they was unaware that 'rose' could function as the past tense of 'rise'. The same participants who misread 'rose' in Test 1 showed analogous patterns in Test 2.

Contextual and Semantic Strategy Outcomes

Afflicted in Test 1 produced 88.5% accuracy, and *somnolent* in Test 2 produced 96.2%. For *somnolent*, most correct responses relied on a single adjacent word: 'sleepiness'. The text reads: "There was a sylvan attitude of somnolent sleepiness pervading that section...". P2 stated explicitly: "'Sleepiness', may be because that word, I think, is a clue for the answer". For *afflicted*,

incorrect participants recognized the sentence's negative emotional tone. P2 explained their choice by noting, "The sentence refers to something negative". This led them to select the option "Confused by a sudden situation" instead of the correct "Suffering from a medical condition". The sentence describes breaking the news of a husband's death.

Syntactic Strategy Outcomes

Forestall in Test 1 produced 96.2% accuracy, and *acceded* in Test 2 produced 88.5%. The open-ended responses showed that participants grabbed onto a single visible word nearby, like "hastened" for *forestall* or "begged me tearfully" for *acceded*. Three participants answered *acceded* incorrectly. The text reads: "Bill begged me tearfully to let him go. So, to relieve Bill, I acceded". Both P1 and P2 thought Bill was the subject of acceded. P1 explained their logic: "I was thinking of Bill. Because he was begging so tearfully, I thought he was the one rushing to leave" (original: "Aku mikir Bill karena begged tearfully, I think dia berusaha dengan tergesa-gesa mau pergi"). P2 echoed this: "It was Bill who left, because it also said 'let him go'" (original: "Bill yang pergi kak, karena disitu juga ada kalimat 'let him go'"). P4 stated, "I imagined it was about a family conflict (whether Bill had a conflict with his wife or something)".

World Knowledge Strategy Outcomes

Elusive in Test 1 produced 96.2% accuracy. *Diatribe* in Test 2 produced 92.3% accuracy, with two students answering incorrectly. These two incorrect responses were split evenly between Option A (A type of modern farming equipment) and Option B (A large financial donation).

Table 2. Distribution of Answer Choices for Diatribe (Test 2, n=26)

Option	Answer Choice	n (%)
A	A type of modern farming equipment	1 (3.8%)
B	A large financial donation	1 (3.8%)
C	A forceful and bitter verbal/written attack	24 (92.3%)
D	A positive review of a business	0 (0.0%)

The sentence containing *diatribe* reads: "We knew that Summit couldn't get after us with anything stronger than constables and, maybe, some lackadaisical bloodhounds and a diatribe or two in the Weekly Farmers' Budget". P2 stated: "Yes, because there was the word 'Budget', so I thought it was related to financial matters" (original: "Betul kak, karena ada kata 'Budget' jadi dipikirnya ada hubungannya sama financial"). When asked whether their answer would have changed without the words 'Budget' and 'Farmers', they replied, "Probably my answer would be different". P5, who answered correctly, explained, "I was slightly misled toward option B because that's the most common meaning, but I immediately noticed that 'Weekly Farmers' Budget' was a newspaper name. Then I chose option C" (original: "Iya memang agak terkecoh ke opsi B karena itu arti yang paling umum didenger, tapi kemarin langsung ngeh kalau ada kata kunci weekly farmers yang artinya nama surat kabar, jadi milih opsi c").

Interlingual Transfer Strategy Outcomes

Intelligence in Test 1 produced 69.2% accuracy, with eight of twenty-six students answering incorrectly. *Reconnoitre* in Test 2 produced 96.2% accuracy, with only one incorrect answer. For *intelligence*, the text uses it to mean news or information received at a newspaper office. Six of the eight participants who answered incorrectly all selected "High mental ability or brain power". For reconnoitre, participants recognized the word's similarity to "rekon" or "observasi". P4 described their thought process: "The Indonesian word 'observasi' came to mind, and I combined that with the context" (original: "Muncul kata Indonesia 'observasi' di pikiran saya, dan saya kombinasikan dengan konteks"). P5 illustrated: "I immediately rejected the meaning of 'smartness' because of the keyword 'newspaper office,' which automatically connected me to 'intel' or news information" (original: "Waktu itu saya langsung nolak arti kecerdasan kak karena ada kata kunci newspaper office yang membuat saya otomatis menghubungkan dengan istilah intel atau informasi berita"). P3 admitted: "I didn't pay much attention to the other phrases; I just guessed directly from the word intelligence" (original: "Iya kak, saya nggak terlalu merhatiin frasa lain, jadi langsung nebak lewat intelligence nya"). For *reconnoitre*, P3 explained, "Before even reaching the word reconnoitre, I was already thinking he would rest, because of 'I told Bill to wait' and the mountain setting" (original: "Iya, sebelum kata reconnoitre kak lebih kepikiran karena frasa 'I told bill to wait'").

Monitoring Strategy Self-Reports

P2 openly admitted their own monitoring failure: "At that time I didn't really re-check the whole sentence, I trusted the word too quickly" (original: "Nggak terlalu ngecek ulang seluruh kalimatnya, terlalu cepat percaya sama kata yang dipikirnya ada hubungannya sama jawabannya"). P5 consistently applied a three-step process: forming an initial impression, noticing a potential mismatch, and correcting themself using the context. They processed a stretch of text as a single meaningful phrase, reading 'rose and fell' instead of just 'rose,' and "Weekly Farmers' Budget" instead of 'Budget'. P5 warned, "It can be dangerous if we're too confident without looking at the sentence context".

Discussion

Across all ten target words, one pattern held: high accuracy did not always mean a clean inferencing process. The findings below evaluate these patterns organized by Nassaji's (2003) five strategy categories, with a cross-cutting observation about monitoring strategy discussed at the end.

Morphological Strategy: False Lexical Anchor Overrides Word-Level Analysis

The accuracy gap between these two morphological items points to a more fundamental issue in how students process literary text. The difficulty with *tumultuously* was not a failure of

morphological analysis. Rather, students never reached the target word. The data show that the literary polysemy of 'rose', functioning simultaneously as a common noun and as the past tense of rise, provided a false contextual anchor that was more immediately accessible than the morphological structure of the target word (Carrol et al., 2018). A prerequisite for successful inferencing is selecting the appropriate knowledge source in the first place Nassaji (2003); in this case, participants selected the wrong source entirely. Orthographic complexity in literary vocabulary can apparently discourage strategy activation before it even starts. The interaction between literary polysemy and a vocabulary gap thus produced a failure that neither factor alone would predict. This finding aligns with Nassaji's (2004) demonstration that depth of vocabulary knowledge determines not only what strategies learners use, but whether those strategies succeed. The fact that the same participants who misread 'rose' showed analogous false anchor patterns in Test 2 confirms that the root problem lies in strategy application, not just in unfamiliarity with the text.

Contextual and Semantic Strategy: Emotional Register Without Semantic Precision

The open-ended responses reveal that the students' high scores actually hide a flawed inferencing process. Because a near-synonym appeared immediately after the target word *somnolent*, actual inferencing was barely needed. The adjacent word acted as a direct translation. High accuracy on this item, therefore, points to simple vocabulary recognition, not contextual strategy skill. Without that adjacent clue, the students likely would have failed. The errors on *afflicted* show a different problem. Because the sentence describes breaking the news of a husband's death, participants easily picked up on the grief and delicacy. However, they could not extract the specific medical domain from that emotional atmosphere alone. Alkhammash (2022) notes a similar phenomenon: non-native speakers reading figurative language often grasp the general evaluative tone without understanding the precise lexical definition. This limitation clearly extends to literary prose. Literature frequently relies on feeling and implication rather than strict denotation. As a result, EFL learners walk away with a correct emotional read of the scene, but no actual vocabulary comprehension (Carrol et al., 2018).

Syntactic Strategy: Grammatical Subject Misidentification and Narrative Substitution

The high numbers hide what was actually happening during the reading process. If that shortcut word pointed in the right direction, the strategy worked. If it was ambiguous, the strategy collapsed. The interviews point to two different kinds of breakdowns. The first is a simple grammatical mix-up. Participants understood the emotional tension perfectly, but attached the action to the wrong person. This was a pronoun tracking error, not a failure of context. The second breakdown is narrative substitution. Tracking syntactic relationships across multiple clauses is inherently demanding. The data here show what can happen when that demand becomes too great: rather than working through the text, participants built a separate narrative and used that instead. Authentic literature makes this escape route very tempting. Its shifting pronouns, complex clause structures, and character-driven context create conditions where narrative substitution is particularly likely to occur.

World Knowledge Strategy: Literary False Clue as Systematic Distractor

Rather than random guessing, these specific errors point directly to how the surrounding textual clues bifurcated the students' interpretations and actively misled them. A reader familiar with early twentieth-century American print culture would recognize this proper noun immediately. However, an Indonesian EFL student lacking that specific cultural schema reads the individual words literally. The word farmers triggers an agricultural schema, while the word budget pulls their interpretation straight toward a financial context. The response shows how a culturally embedded literary clue actively misled them. The difference is that a successful reader recognized the phrase as a compound proper noun instead of three separate content words. This is a gap in cultural schema, not vocabulary. Recognizing the naming conventions of small-town American newspapers requires specific background knowledge. EFL students cannot realistically be expected to possess this without explicit instruction.

Interlingual Transfer Strategy: Productive Transfer, Misleading Transfer, and Narrative Blocking

This 27 percentage point gap between two words in the same strategy category is the most theoretically significant difference in the quantitative data, and it requires explanation. The open-ended responses showed that nearly all participants relied on L1 transfer for both words. The difference was the direction of that transfer. For *intelligence*, the transfer was highly misleading. The consistency of this error reflects how strongly the L1 meaning of *intelligence* overrode the contextual evidence available in the sentence. For *reconnoitre*, however, the transfer worked. This L1 bridge led them straight to the correct meaning. The success of interlingual transfer did not depend on the transfer itself. It depended entirely on whether the student verified their guess against the context. P3 accepted the L1 transfer blindly, without any contextual cross-checking. Furthermore, the narrative context alone activated a predictive schema: the character climbed the mountain, reached the top, and rested. This schema filled the meaning slot before any word-level analysis could even begin. This pattern can be described as narrative prediction: the reader's engagement with story logic fills a lexical gap before any word-level strategy is applied. Nassaji (2003) does include discourse-level inference as a valid knowledge source, but his framework treats it as one option among five rather than as a force that can shut the others down entirely. What P3's case illustrates is a more extreme version of discourse-level processing, one where narrative momentum overtakes rather than supports inferencing. Whether this constitutes a genuinely new category or an underexplored edge case of Nassaji's existing framework is a question that warrants further investigation with a larger sample. While contextual information generally supports reading comprehension (Grabe & Stoller, 2019), P3's case demonstrates that narrative engagement can work against lexical inferencing in literary texts. The richness of a literary story can dominate the reader's attention, filling lexical gaps prematurely. Alahmadi and Foltz (2020) found that lexical inferencing yields vocabulary gains only when learners actively engage with the target word. P3's case shows that this active engagement is never met if the narrative context has already answered the question.

Monitoring Strategy as a Cross-Cutting Determinant

Monitoring strategy was the most consistent factor separating successful from unsuccessful inferencing across all five strategy categories. It was not the only factor at play. Vocabulary gaps still trapped P1, and a lack of cultural background misled P2, regardless of how well they monitored their reading. However, monitoring was the single variable traceable across all five strategy categories in the interview data. Nassaji (2003) describes monitoring as a metacognitive process where readers evaluate their own guesses, check provisional meanings against the wider context, and revise when things do not add up. Similarly, Rastegar et al. (2017) link metacognitive strategy use directly to reading comprehension achievement. The current data illustrates exactly what this looks like in practice, and what happens when it goes missing. Their self-report highlights the exact mechanism of failure. A single word triggered a plausible association, and they accepted that meaning without verifying it against the rest of the sentence. This exact sequence, rapid association leading to premature closure, drove P1's and P2's mistake with 'rose', P3's narrative prediction, and P4's invented storyline. Monitoring would have inserted a crucial pause between the first impression and the final answer. For these students, that pause never happened. Nassaji (2003) frames monitoring as checking provisional meanings. P5 went further. They questioned the reliability of their own cognitive strategy before accepting its output, demonstrating the most sophisticated metacognitive awareness in the group. A final, crucial difference lies in the size of the reading unit. The present study refers to this specific behavior as reading in units. Processing 'rose and fell' as a single unit instantly activates the meaning of breathing, leaving no room for the flower association to emerge. Conversely, reading "Budget" in isolation practically guarantees a financial misinterpretation. Nassaji's taxonomy does not explicitly categorize this reading in units behavior. It acts more like a prerequisite cognitive habit. It determines whether a student starts applying a strategy from a solid anchor or a false one. Finally, the gap between P5 and the rest of the class was not just about which strategy they chose. It was about how much of the text they actually read before that strategy was even triggered.

Theoretical Contributions and Novelty

While previous studies have established lexical inferencing as a core reading strategy and noted the benefits of familiar local culture in reading comprehension, existing literature has largely treated reading texts as interchangeable. This tendency heavily underrepresents the specific cognitive demands of authentic literary prose (Yang et al., 2023). Previous researchers have not mapped the precise cognitive breakdown points within established inferencing frameworks when readers encounter literary polysemy, nor have they fully explained how narrative momentum or unfamiliar cultural allusions systematically override linguistic analysis. Addressing this gap, the present research found that high accuracy in inferencing often masks deeply flawed cognitive processes. Specifically, this study identified several key breakdown points: literary polysemy creates false lexical anchors that override morphological analysis,

while complex syntax leads to narrative substitution where readers invent alternative storylines rather than parsing the sentence. Furthermore, readers often over-rely on emotional registers at the expense of semantic precision, and unfamiliar cultural clues act as systematic distractors. Crucially, the data revealed that strong engagement with the plot leads to narrative prediction, a process that prematurely fills lexical gaps and blocks word-level strategy application. Throughout these processes, the presence or absence of metacognitive monitoring and the habit of "reading in units" determine whether an inferencing strategy succeeds or fails.

Consequently, the novelty of this research lies in its shift from evaluating general inferencing outcomes, such as simple success or failure rates, to mapping the specific cognitive breakdown pathways triggered by authentic literary features. This study introduces and details a concept not fully accounted for in existing reading frameworks: "Narrative Prediction," where story logic completely overtakes linguistic analysis. This novel pathway provides a more granular understanding of how lexical inferencing operates and breaks down within the complex environment of authentic literature, offering a necessary theoretical expansion to existing inferencing frameworks that currently assume a more balanced interaction between contextual and linguistic knowledge.

Conclusion

This study investigated the difficulties encountered by the fourth-semester EFL students when applying Nassaji's (2003) five lexical inferencing strategies in authentic literary texts, as well as the linguistic and non-linguistic factors underlying those difficulties. Findings show that all five strategy categories were activated, but each broke down at a different point. Morphological strategy was preempted before the target word was even reached because literary polysemy in a nearby word drew students' attention away. Contextual strategy captured the emotional register of a sentence but could not convert that into semantic precision. Syntactic strategy failed through two routes: misidentification of the grammatical subject, and the wholesale replacement of the text with an imagined narrative. World knowledge strategy was actively misled by culturally embedded words functioning as false clues. Interlingual transfer produced opposite outcomes in the same category, succeeding when students verified their L1 association against the context and failing when they did not. An additional pattern, narrative prediction as a blocker of word-level analysis, was identified as a process not fully accounted for by Nassaji's existing categories, in which narrative engagement fills lexical gaps before any linguistic strategy has a chance to operate. The false anchor pattern, in particular, appeared consistently across both textual conditions for the same participants, suggesting that at least this difficulty is rooted in features of literary language rather than in unfamiliarity with a specific text.

By identifying phenomena such as narrative prediction and the critical role of reading-in-units behavior, this research moves the body of scientific knowledge forward by expanding established inferencing frameworks to account for the unique cognitive demands of literary prose. However, this study is not without limitations. The data collection relied on unproctored online questionnaires, meaning the use of external resources could not be definitively ruled out.

Furthermore, the findings are based on the responses of a specific group of fourth-semester students reading two selected short stories, which may limit the broad generalization of the cognitive patterns observed.

These findings carry direct practical implications for EFL reading instruction. Rather than merely teaching vocabulary strategies in isolation, teachers should explicitly train students to read in phrasal chunks (e.g., instructing students to process rose and fell as a single unit rather than fixating on rose). Furthermore, instructors should deliberately design exercises that expose students to literary false clues, training them to delay semantic closure until they have verified their initial guesses against the broader sentence context. Future research might specifically investigate whether narrative prediction as a blocker of word-level analysis also operates in other literary genres, particularly in first-person narratives, where character psychology dominates the prose, and whether explicit instruction in reading-in-units behavior can reduce false anchor selection in literary inferencing tasks.

References

Alahmadi, A., & Foltz, A. (2020). Exploring the effect of lexical inferencing and dictionary consultation on undergraduate EFL students' vocabulary acquisition. *PLoS ONE*, *15*(7 July), 1–25. https://doi.org/10.1371/journal.pone.0236798

Ali, M. A. (2020). Investigation of Vocabulary Learning Strategies to Identify Word Meanings for Saudi EFL Students in Reading Context. *Arab World English Journal*, *11*(3), 149–169. https://doi.org/10.24093/awej/vol11no3.9

Alkhammash, R. (2022). Processing figurative language: Evidence from native and non-native speakers of English. *Frontiers in Psychology*, *13*(November), 1–13. https://doi.org/10.3389/fpsyg.2022.1057662

Bobkina, J., & Stefanova, S. (2016). Literature and critical literacy pedagogy in the EFL classroom: Towards a model of teaching critical thinking skills. *Studies in Second Language Learning and Teaching*, *6*(4), 677–696. https://doi.org/10.14746/ssllt.2016.6.4.6

Bunparit, C., & Chinokul, S. (2018). Different Proficiency EFL Learners' Employment of Lexical Inferencing Strategies in Reading Academic Texts. *Journal of Education*, *19*(2), 103–120. https://ejournals.swu.ac.th/index.php/jedu/article/view/11709/9713

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: complex or simple? Research case examples. *Journal of Research in Nursing*, *25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Carrol, G., Littlemore, J., & Gillon Dowens, M. (2018). Of false friends and familiar foes: Comparing native and non-native understanding of figurative phrases. *Lingua*, *204*, 21–44. https://doi.org/10.1016/j.lingua.2017.11.001

Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (4th ed.). SAGE Publications.

Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). CONTEXTUAL WORD LEARNING DURING READING IN A SECOND LANGUAGE. *Studies in Second Language Acquisition*, *40*(2), 341–366. https://doi.org/10.1017/S0272263117000109

Fauzi, I., Asi, N., Saman, T. N., Karani, E., Luardini, M. A., & Ngabut, M. N. (2024). Examining Teaching Reading Using Local Culture-Based Pictorial Narrative Texts: A Case Study to Secondary School Students. *Jurnal Pendidikan Edutama*, *11*(1), 36. https://doi.org/10.30734/jpe.v11i1.3987

Grabe, W., & Stoller, F. L. (2019). Teaching and researching reading, third edition. In *Teaching and Researching Reading*. Routledge. https://doi.org/10.4324/9781315726274

Hassanzadeh, Z., Tamjid, N. H., Ahangari, S., Dios, J. De, & Agudo, M. (2019). The effect of lexical inference strategy instruction on Iranian EFL learners ' vocabulary depth and breadth. *Cogent Education*, *00*(00), 1–16. https://doi.org/10.1080/2331186X.2019.1614750

Kaufmann, K., & Peil, C. (2020). The mobile instant messaging interview (MIMI): Using WhatsApp to enhance self-reporting and explore media usage in situ. *Mobile Media and Communication*, *8*(2), 229–246. https://doi.org/10.1177/2050157919852392

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of Experimental Child Psychology*, *160*, 1–20. https://doi.org/10.1016/j.jecp.2017.02.015

Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative Data Analysis: A Methods Sourcebook* (4th ed.). SAGE Publications.

Nassaji, H. (2003). L2 Vocabulary Learning from Context: Strategies, Knowledge Sources, and Their Relationship with Success in L2 Lexical Inferencing. *TESOL Quarterly*, *37*(4), 645. https://doi.org/10.2307/3588216

Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *Canadian Modern Language Review*, *61*(1), 107–124. https://doi.org/10.3138/cmlr.61.1.107

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.

Pellicer-Sánchez, A. (2016). INCIDENTAL L2 VOCABULARY ACQUISITION FROM AND WHILE READING. *Studies in Second Language Acquisition*, *38*(1), 97–130. https://doi.org/10.1017/S0272263115000224

Peters, E., & Webb, S. (2018). INCIDENTAL VOCABULARY ACQUISITION THROUGH VIEWING L2 TELEVISION and FACTORS THAT AFFECT LEARNING. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Rastegar, M., Mehrabi Kermani, E., & Khabir, M. (2017). The Relationship between Metacognitive Reading Strategies Use and Reading Comprehension Achievement of EFL Learners. *Open Journal of Modern Linguistics*, *07*(02), 65–74. https://doi.org/10.4236/ojml.2017.72006

Rose, H., McKinley, J., & Briggs Baffoe-Djan, J. (2019). *Data Collection Research Methods in Applied Linguistics* (1st ed.). Bloomsbury Academic.

Rutamornchai, N., & Tepsuriwong, S. (2022). When Lexical Inferencing Failed: An Investigation into Strategy Use and Factors Underpinning Less Successful Lexical Inferencing. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 124–155.

Yang, H., Fan, L., & Yin, H. (2023). Knowledge mapping of the research on lexical inferencing: A bibliometric analysis. In *Frontiers in Psychology* (Vol. 14, Number January, pp. 1–13). https://doi.org/10.3389/fpsyg.2023.1101241

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing*, *25*(5), 1195–1216. https://doi.org/10.1007/s11145-011-9313-z