FROM CRAMMING TO DURABLE VOCABULARY LEARNING: A RETRIEVAL-AND-SPACING FRAMEWORK FOR VIETNAMESE UPPER-SECONDARY EFL CLASSROOMS

Nguyen Anh Tuan[1]
[1]*Head of Upper-Secondary English Department, Pétrus Ký Primary - High School, Ho Chi Minh City, Vietnam*
Email address: anhtuan@pek.edu.vn
Corresponding author: Nguyen Anh Tuan
https://doi.org/10.51453/3093-3706/2026/1501

ARTICLE INFO

Received: 15/06/2026
Revised: 20/07/2026
Published: 28/08/2026

KEYWORDS

Retrieval practice;
Spaced practice;
Vocabulary retention;
Upper-secondary education.

ABSTRACT

Vocabulary is central to English-as-a-foreign-language (EFL) learning, yet classroom instruction can create a gap between initial word exposure and durable access to word knowledge. This integrative review synthesizes research on retrieval practice, spaced practice, cumulative testing, and vocabulary task format, and relates the evidence to Vietnamese upper-secondary EFL conditions. Twenty-one core publications were selected for relevance to Vietnamese vocabulary education, retrieval-based learning, spacing, and classroom implementation. The synthesis identifies four design principles: require successful effortful retrieval rather than repeated restudy; distribute retrieval opportunities across lessons; align retrieval mode with the target form of knowledge; and move from form-meaning recall to contextualized production. On this basis, the article proposes a three-stage Retrieve-Revisit-Recontextualize (3R) framework. The framework embeds short, low-stakes retrieval episodes into ordinary lessons, uses cumulative revisiting over increasing intervals, and finishes with productive reuse in sentences, speaking, or short writing. The framework is intended as a feasible pedagogical design rather than a tested intervention. It offers a research-informed basis for future classroom trials in Vietnamese upper-secondary schools.

TỪ HỌC DỒN ĐẾN GHI NHỚ TỪ VỰNG BỀN VỮNG: KHUNG LUYỆN TRUY HỒI KẾT HỢP GIÃN CÁCH CHO LỚP TIẾNG ANH THPT VIỆT NAM

Nguyễn Anh Tuấn[1]
[1]*Trưởng bộ môn Tiếng Anh cấp Trung học phổ thông, Trường Tiểu học - Trung học phổ thông Pétrus Ký, Thành phố Hồ Chí Minh, Việt Nam*
Địa chỉ Email: anhtuan@pek.edu.vn
Tác giả liên hệ: Nguyễn Anh Tuấn
https://doi.org/10.51453/3093-3706/2026/1501

THÔNG TIN BÀI BÁO

Ngày nhận bài: 15/06/2026
Ngày hoàn thiện: 20/07/2026
Ngày đăng: 28/08/2026

TỪ KHOÁ

Luyện truy hồi;
Luyện tập giãn cách;
Ghi nhớ từ vựng;
Trung học phổ thông

TÓM TẮT

Từ vựng giữ vai trò trung tâm trong học tiếng Anh, nhưng dạy học trên lớp có thể tạo khoảng cách giữa tiếp xúc ban đầu và khả năng truy xuất từ vựng bền vững. Bài tổng quan tích hợp các nghiên cứu về luyện truy hồi, giãn cách, kiểm tra tích lũy và dạng nhiệm vụ từ vựng, đồng thời liên hệ bằng chứng với bối cảnh tiếng Anh THPT Việt Nam. Hai mươi mốt công trình được lựa chọn theo mức độ liên quan đến giáo dục từ vựng tại Việt Nam, học tập dựa trên truy hồi, giãn cách và triển khai trong lớp học. Tổng hợp chỉ ra bốn nguyên tắc: ưu tiên truy hồi có nỗ lực thay cho học lại thụ động; phân bố cơ hội truy hồi qua nhiều tiết học; lựa chọn dạng truy hồi phù hợp với loại kiến thức cần hình thành; và chuyển từ nhớ nghĩa sang sử dụng từ trong ngữ cảnh. Từ đó, bài viết đề xuất khung ba bước Truy hồi-Tái gặp-Tái ngữ cảnh hóa (3R), hướng đến các hoạt động ngắn, ít áp lực, có tính tích lũy và khả thi trong lớp THPT.

1. INTRODUCTION

Vocabulary learning is not simply a matter of meeting a word once and recognizing its meaning. For learners to use vocabulary in reading, listening, speaking, and writing, lexical knowledge must remain accessible after the lesson in which it was first presented. This distinction between initial exposure and durable retention is particularly important in foreign-language contexts, where classroom time may provide the main regular source of English input and practice. A review of vocabulary learning, teaching, and testing in Vietnam concluded that Vietnamese EFL learners often show limited knowledge of both single words and formulaic language and that vocabulary pedagogy would benefit from a more principled approach to repetition, attention, and strategy training (D. V. Vu & Peters, 2021).

The school context makes the issue more specific. A recent study of Vietnamese upper-secondary learners reported generally positive attitudes toward learning word form, meaning, and use (D. H. Tran, 2026). Positive attitudes, however, do not by themselves determine whether lexical knowledge will still be available days or weeks later. At the same time, analyses of Vietnamese upper-secondary textbooks indicate substantial lexical demands and show that textbook series vary in the quantity and distribution of vocabulary learners encounter (Lien et al., 2024). Corpus-based examination of a Vietnamese high-school EFL textbook similarly highlights the importance of lexical coverage and high-frequency vocabulary for making input accessible (N. T. M. Le & Dinh, 2022). These findings suggest that the instructional problem is not a lack of vocabulary content. Rather, teachers need a manageable way to help students retain and retrieve important words after initial teaching.

One common response to the pressure of vocabulary coverage is massed review: learners reread a list, copy definitions, repeat a set of exercises, or review intensively immediately before a quiz. Such activity may produce a feeling of familiarity and can improve immediate performance, but familiarity is not identical to retrievability. Research on the testing effect has shown that attempting to retrieve previously learned information can itself strengthen subsequent retention. In a landmark foreign-word learning study, repeated testing after successful learning produced substantially better delayed recall than continued restudy (Karpicke & Roediger, 2008). A meta-analysis across testing-effect research later confirmed a robust advantage of testing over restudy while showing that effect size depends on test format, retrieval success, and other design conditions (Rowland, 2014).

Second-language vocabulary research has subsequently refined the question from whether retrieval helps to how retrieval should be scheduled. Studies have examined the number of retrievals within a session, cumulative testing over a semester, spacing between learning opportunities, activity type, modality, and combinations with other memory supports. The resulting evidence is useful but not automatically classroom-ready. Vietnamese upper-secondary teachers work within fixed lesson sequences, mixed proficiency levels, examination pressure, limited time for individual monitoring, and a need to integrate vocabulary with reading, grammar, speaking, and writing rather than teach vocabulary as a separate course.

This article therefore aims to translate retrieval-and-spacing research into a practical instructional design for Vietnamese upper-secondary EFL classrooms. It addresses two questions: (1) What classroom design principles can be derived from research on retrieval practice and spaced practice for L2 vocabulary learning? and (2) How can those principles be organized into a low-burden framework suitable for ordinary Vietnamese upper-secondary English lessons? The contribution is conceptual and pedagogical. No claim is made that the proposed framework has already been validated in Vietnamese schools; instead, it offers an evidence-informed model that can be tested in future empirical studies. The next section reviews existing evidence and the research gap; subsequent sections explain the review method, derive the 3R framework, and discuss its implications and limits.

2. LITERATURE REVIEW

2.1. Evidence from Vietnamese upper-secondary education

Existing research in Vietnam highlights several complementary aspects of the problem. D. V. Vu and Peters (2021) review vocabulary knowledge, teaching, and testing in Vietnamese English-as-a-foreign-language education. Lien et al. (2024) and N. T. M. Le and Dinh (2022) document textbook vocabulary demands and lexical coverage, whereas D. H. Tran (2026) examines upper-secondary learners' attitudes toward vocabulary learning. These studies clarify what students encounter and value, but do not directly test the durability of a cumulative retrieval schedule.

Vietnamese school studies provide useful but limited implementation precedents. In an action-research study involving 30 Grade 11 students in Thai Nguyen, M. D. Tran and

Nguyen (2022) report improved vocabulary retention and favourable views after Quizlet use. Huyen and Nghi (2026) examine Kahoot!-mediated instruction with two Grade 11 groups (27 and 35 students), reporting vocabulary gains and positive attitudes. D. T. Tran and Le (2023) study Quizlet with 300 Grade 8 students and 15 teachers in Binh Duong. These interventions involve multiple platform features and do not separately test the three 3R stages; Grade 8 results also require caution when applied to upper-secondary learners.

Evidence on teaching conditions adds a distinct perspective: T. V. Le et al. (2024) study high-school teachers' views on teaching reform and assessment, Nguyen (2024) examines teachers' roles during tasks, and H. H. Vu (2025) discusses how teachers perceive changes in the national upper-secondary English examination. These sources strengthen the practical rationale for brief, curriculum-compatible routines, but they cannot establish vocabulary retention gains.

2.2. International evidence on retrieval and spacing

Research on retrieval practice distinguishes a successful attempt to recall from passive re-exposure (Karpicke & Roediger, 2008; Rowland, 2014). For L2 vocabulary, Nakata (2017) examines retrieval frequency, Nakata et al. (2021) investigate cumulative classroom quizzes, and Kim and Webb (2022) synthesize evidence for spaced second-language practice. More recent studies address activity type and relearning intervals (Kim & Webb, 2023; Nakata et al., 2023), spoken and written test modality (Uchihara, 2023), and mnemonic support for difficult retrieval (Qu et al., 2024). The findings motivate design choices but do not imply a single optimal interval or guarantee transfer from word-pair recall to communicative production.

2.3. Research gap

Although international research establishes relevant learning mechanisms, the Vietnamese sources reviewed here focus mainly on textbook demands, learner attitudes, digital vocabulary interventions, teacher practice, or examination conditions. Within this selected literature, no study directly tests an integrated, low-technology Retrieve-Revisit-Recontextualize sequence with the same target items, comparable review time, and delayed receptive and productive outcomes in Vietnamese upper-secondary classes. This is a gap identified within the reviewed corpus, not a claim that no such work exists elsewhere. The present article responds by formulating a transparent, classroom-manageable design for later empirical evaluation.

3. RESEARCH METHODS

3.1. Review design and source identification

The article used an integrative review approach rather than a systematic meta-analysis. The purpose was to connect findings from cognitive psychology, second-language vocabulary research, and recent Vietnamese EFL literature in order to derive classroom design principles. Searches and source checks were conducted through scholarly web indexes, publisher databases, ERIC, PubMed-indexed records, and journal websites. Search combinations included retrieval practice, testing effect, spaced practice, distributed retrieval, second-language vocabulary, EFL vocabulary, digital flashcards, Vietnamese vocabulary learning,

and upper-secondary English. A targeted supplementary search was also undertaken to identify additional Vietnamese classroom vocabulary and teaching-context studies.

3.2. Eligibility, exclusion, and screening record

Sources were retained when they met at least one of four criteria: they provided foundational evidence for retrieval practice; reported a meta-analysis or controlled study of spacing in L2 learning; investigated retrieval, spacing, cumulative quizzes, or task modality specifically for L2 vocabulary; or supplied recent evidence about vocabulary learning and curricular conditions in Vietnam. Priority was given to peer-reviewed sources with DOI information and to studies whose methods and outcomes were sufficiently described to support pedagogical interpretation. Twenty-one publications formed the reviewed evidence base. The review was intentionally selective, because its goal was design synthesis rather than an exhaustive estimate of effect size.

Sources were excluded from the analytical synthesis when they were off-topic, lacked traceable publication or methodological details, offered only anecdotal or promotional claims, or focused on an education level without a relevant contribution to the proposed upper-secondary design.

A documented shortlist of 25 candidate publications was assessed using titles, abstracts, publication information, and available methods. Of these, 21 met the inclusion criteria; four were excluded because their participants were outside the target educational level and the studies offered no distinct insight for the upper-secondary framework. The figures refer to the inspected shortlist rather than the full yield of database searches.

3.3. Extraction, source-quality appraisal, and synthesis

For each source, the review extracted the learning context, type of vocabulary task, retrieval or spacing manipulation, outcome timing, and practical implication. Findings were then grouped into four themes: retrieval as a learning event; spacing and cumulative revisiting; alignment between retrieval mode and target knowledge; and movement from decontextualized recall toward productive use. These themes were compared with the constraints reported or implied in Vietnamese secondary vocabulary research. The final framework was constructed only from principles that could be implemented without specialized software or substantial additional class time.

A qualitative quality appraisal considered (a) whether source metadata and methodological information were traceable, (b) whether study design supported causal, descriptive, or contextual conclusions, (c) whether vocabulary outcomes were measured directly and after a delay, and (d) the applicability of participants and tasks to Vietnamese upper-secondary education. Experimental and meta-analytic results were used to justify cognitive learning mechanisms; school-based action studies informed classroom feasibility; surveys, textbook analyses, and teacher reports informed context, not direct 3R effects. At the individual-study level, Rowland (2014) aggregates different testing formats, while the 30-student action study (M. D. Tran & Nguyen, 2022) cannot isolate Quizlet features from

retrieval effects. Appraisal was completed by one reviewer without a numerical risk-of-bias score or independent double-rating, and this constraint limits the strength of the synthesis.

Table 1 summarizes the evidence-to-design appraisal. It records both the contribution and the principal inference limits of the different types of sources. No effect sizes were pooled because the included studies differ substantially in samples, tasks, outcomes, and retention intervals.

Table 1. Qualitative appraisal of the main evidence groups

Evidence group	**Contribution to 3R**	**Quality and transferability limitation**
Retrieval experiments and reviews (Karpicke & Roediger, 2008) (Nakata, 2017) (Rowland, 2014)	Support genuine retrieval attempts, feedback, and attention to practice efficiency.	Stronger learning-mechanism evidence; many participants and tasks are not Vietnamese THPT contexts.
L2 spacing and modality studies (Kim & Webb, 2022) (Kim & Webb, 2023) (Nakata et al., 2021) (Nakata et al., 2023) (Uchihara, 2023)	Support cumulative revisiting, variable spacing, and matching response mode to target knowledge.	Controlled studies and synthesis; different delays and outcomes do not establish one ideal 3R timetable.
Vietnamese school vocabulary interventions (Huyen & Nghi, 2026) (D. T. Tran & Le, 2023) (M. D. Tran & Nguyen, 2022)	Demonstrate feasibility of monitored vocabulary practice in local schools.	Grade 8/11 groups and multi-component apps; no isolated test of the combined low-tech 3R model.
Vietnamese textbook, attitude, and teacher studies (N. T. M. Le & Dinh, 2022) (T. V. Le et al., 2024) (Lien et al., 2024) (Nguyen, 2024) (D. H. Tran, 2026) (D. V. Vu & Peters, 2021) (H. H. Vu, 2025)	Inform word selection, lesson pacing, feedback, and assessment constraints.	Descriptive/corpus/survey/observation evidence; cannot establish causal vocabulary-retention effects.

4. RESULTS AND DISCUSSION

4.1. Retention should be treated as a separate instructional goal

A first implication of the literature is that teaching a word and retaining a word should be treated as related but distinct instructional problems. Vietnamese learners may meet large amounts of vocabulary through textbooks, reading passages, teacher explanations, and test-preparation materials (N. T. M. Le & Dinh, 2022; Lien et al., 2024; D. V. Vu & Peters, 2021). Yet the number of encounters is pedagogically useful only when some of those encounters require learners to reconstruct knowledge rather than merely see it again. This helps explain why a student can say that a word looks familiar but still fail to produce its meaning, form, collocation, or appropriate use under test or communicative pressure.

The relevant unit of planning is therefore not only the vocabulary list for a lesson but also the future retrieval opportunities attached to that list. If ten target words are introduced on Monday and disappear from instruction after Tuesday, the lesson sequence optimizes coverage rather than retention. By contrast, a retention-oriented sequence intentionally brings a smaller set of high-value items back after forgetting has begun. This principle is compatible with the Vietnamese context because it does not require replacing the textbook. It changes what happens after first presentation: selected items are recycled across later warm-ups, exit tickets, short quizzes, sentence-completion tasks, and speaking or writing prompts.

4.2. Retrieval is more than assessment

Retrieval practice refers to attempting to bring previously learned information to mind without simply restudying the answer. Its instructional value is easy to underestimate because the activity often resembles a test. However, research distinguishes retrieval used to measure learning from retrieval used to produce learning. Karpicke and Roediger (2008) demonstrated with foreign-language word pairs that repeated retrieval after an item had initially been learned supported delayed retention more effectively than additional study of the same item. Rowland (2014) likewise found a reliable testing effect across a broad research base, with stronger benefits under conditions that require more effortful recall rather than passive recognition.

For L2 vocabulary, the practical implication is that retrieval should be built into the learning phase rather than reserved for a graded test. Nakata (2017) examined one, three, five, or seven retrievals of English-Japanese word pairs and found that more within-session retrievals could improve delayed performance, although efficiency changed when time-on-task was considered. This qualification is important for schools: more retrieval is not automatically better if it consumes disproportionate lesson time. The aim should be sufficient successful retrieval distributed intelligently, not endless drilling.

Low-stakes design also matters. A retrieval attempt should provide diagnostic information to the learner but need not carry marks. Students can answer on mini-whiteboards, paper slips, notebooks, response cards, or a digital form. After an attempt, the correct form should be supplied so that failure does not become repeated rehearsal of an error. The teacher's role shifts from repeatedly explaining the same list to engineering short cycles in which students must produce the answer, check it, and then encounter the item again later.

4.3. Spacing turns retrieval into durable learning

Retrieval becomes more powerful for long-term learning when opportunities are distributed across time. In a semester-long classroom study, Nakata et al. (2021) compared cumulative and noncumulative vocabulary quizzes. Cumulative quizzes repeatedly returned to earlier vocabulary instead of testing only the most recent set, thereby creating distributed retrieval opportunities across the course. This design is attractive for school use because cumulative review can be embedded in existing weekly quizzes without creating a separate intervention.

The broader L2 evidence strongly supports attention to spacing. Kim and Webb (2022) synthesized 48 experiments with 98 effect sizes and found a medium-to-large overall benefit of spaced practice for second-language learning. They also showed that the effect was moderated by factors such as learning target, number of sessions, practice type, activity type, feedback timing, and retention interval. The practical message is not that one universal interval exists. Rather, spacing should be long enough for retrieval to require effort but not so long that most students fail completely.

Recent studies make this point more concrete. Kim and Webb (2023) compared massed and one-day-spaced practice in flashcard and fill-in-the-blank activities. Spaced practice supported better delayed vocabulary performance in both activity types, while the size of the immediate benefit varied with the task. Nakata et al. (2023) further showed that longer within-session spacing during both initial learning and relearning could improve long-term retention and efficiency despite increasing the duration of practice. Together, these findings support a schedule that revisits words across multiple lessons and that adjusts intervals according to retrieval success rather than treating all words identically.

For upper-secondary classrooms, an exact algorithm is unnecessary. A practical starting schedule might revisit a newly taught item at the end of the lesson, in the next lesson, several days later, and again one or two weeks later. Items retrieved accurately on successive occasions can appear less frequently; items that are repeatedly missed can return sooner. This creates a simple form of adaptive spacing without requiring specialized software. The teacher is not attempting to reproduce a laboratory schedule but to preserve the underlying principle: later retrieval should occur after some forgetting, and successful items should not disappear permanently from practice.

4.4. Retrieval format should match the knowledge students need

Vocabulary knowledge is multidimensional. Recognizing that a word has been seen before is not equivalent to recalling its meaning, producing its written form, recognizing its spoken form, selecting an appropriate collocation, or using it accurately in a sentence. This distinction matters because retrieval practice is specific to what learners are asked to retrieve. If every review item is multiple choice, students may become better at recognition without developing the productive access demanded in writing and speaking.

Classroom research on quiz modality provides direct evidence. Uchihara (2023) found that students who prepared for and took spoken-format weekly vocabulary quizzes made larger gains in spoken vocabulary knowledge than students who took written-format quizzes,

whereas written learning did not differ in the same way. The implication is straightforward: the form of retrieval can guide both study behavior and the type of knowledge strengthened. Teachers should therefore vary cues and responses. A target word may first be retrieved from a Vietnamese meaning or English definition, later spelled from an audio cue, then supplied in a sentence gap, and finally used in a short utterance or written sentence.

Retrieval also does not have to work alone. Qu et al. (2024) found that when retrieval opportunities were limited, combining retrieval with a keyword mnemonic improved L2 word retention relative to restudy or restricted retrieval alone. This suggests a useful differentiation principle: students who repeatedly fail to retrieve an item may need stronger encoding support before the next retrieval attempt. A mnemonic, image, word part, first-language link, example sentence, or collocational cue can be used to rebuild the memory trace; retrieval should then follow again rather than ending with the explanation.

4.5. The Retrieve-Revisit-Recontextualize (3R) framework

The evidence can be organized into a three-stage classroom framework: Retrieve, Revisit, and Recontextualize (3R). The framework is deliberately simple. It is not a digital product or a fixed timetable; it is a planning rule that teachers can apply to textbook vocabulary, examination vocabulary, or vocabulary emerging from reading and listening lessons.

Stage 1, Retrieve, begins after initial teaching. Learners close or cover the source and attempt to recall selected target items from cues. The cue should demand the level of knowledge that the teacher wants to strengthen. For receptive meaning, the teacher may present the English word and ask for a brief Vietnamese meaning or English definition. For productive form, the cue can be the meaning and the learner must produce the English word. For orthographic knowledge, an audio cue can require spelling. Feedback follows quickly enough to prevent persistent error, but only after an authentic retrieval attempt has occurred.

Stage 2, Revisit, distributes further retrieval across later lessons. Rather than repeating the whole vocabulary list, the teacher uses a cumulative pool. Newly learned and previously learned items are mixed, which reduces predictability and creates natural spacing. A simple rule is to revisit a new item in the next lesson and then again after several days; items retrieved correctly two or three times can move to longer intervals, while missed items return sooner. Cumulative quizzes described in L2 research (Nakata et al., 2021) provide a classroom precedent for this principle. Spacing research (Kim & Webb, 2022, 2023; Nakata et al., 2023) further indicates that durable learning depends on revisiting across time, not only on the number of repetitions within one sitting.

Stage 3, Recontextualize, prevents the cycle from ending at translation-pair memory. Once students can retrieve form and meaning, they must use the item in a new linguistic context. Tasks can include completing a different sentence, choosing a collocation, transforming a word family member, producing a short spoken response, or writing a sentence related to the current unit. Recontextualization is important because upper-secondary English learning ultimately requires access to words in reading, writing, listening, and speaking, not only success on isolated flashcards. It also creates a bridge between deliberate vocabulary

learning and the meaning-focused input and output recommended for Vietnamese EFL pedagogy (D. V. Vu & Peters, 2021).

Table 2 summarizes the framework and shows how each stage can be implemented with minimal additional time.

Table 2. Retrieve-Revisit-Recontextualize (3R) framework for upper-secondary EFL vocabulary review

Stage	Design principle	Example classroom routine
1. Retrieve	Require a genuine recall attempt before showing the answer; match the cue to the target knowledge.	2-4 minutes after teaching: meaning → English word, English word → brief meaning, audio → spelling, or definition → target word; then provide corrective feedback.
2. Revisit	Return selected items across later lessons; use cumulative and adjustable spacing.	Next lesson: brief warm-up. After 3-4 days: mixed retrieval with older words. After 1-2 weeks: cumulative quiz; missed items return sooner, secure items later.
3. Recontextualize	Move beyond isolated form-meaning recall toward productive and contextualized use.	Use the word in a new sentence, collocation, short spoken response, sentence transformation, or 2-3 sentence writing task linked to the current unit.

4.6. A feasible two-week classroom routine

The 3R framework can operate within a conventional unit without displacing required content. Suppose a teacher identifies 10 to 12 high-value words from a reading or vocabulary lesson. Immediately after teaching, students spend three minutes retrieving the items from mixed cues. At the beginning of the next lesson, five of the items reappear in a two-minute retrieval warm-up. Three or four days later, the set is mixed with vocabulary from the previous unit. At the end of the week, a short cumulative quiz includes both recent and older words. In the following week, the teacher selects several retained items for sentence production, a speaking prompt, or a short paragraph. The total review time can be distributed in small episodes rather than concentrated into a single long vocabulary period.

A crucial management decision is selection. Not every item in a textbook needs equal retrieval practice. Teachers can prioritize high-frequency words, words central to unit objectives, items required for later tasks, useful collocations, and words students repeatedly miss. This reduces workload and prevents the framework from becoming a second syllabus. The Vietnamese textbook analyses cited earlier (N. T. M. Le & Dinh, 2022; Lien et al., 2024) reinforce the need for prioritization because lexical load and coverage vary across materials.

The routine can also be differentiated without producing separate worksheets for every student. After a common retrieval attempt, learners can mark each item with a simple status:

secure, uncertain, or missed. Secure items stay in the cumulative pool but at longer intervals. Uncertain items return in the next review. Missed items receive a brief encoding aid and are retrieved again. Teachers can collect only aggregate information—for example, which five items had the highest error rate—rather than maintaining complex individual profiles. The resulting adaptation is light enough for normal class sizes.

Technology is optional. Digital flashcards and mobile applications can automate scheduling and provide convenient out-of-class practice; recent review work identifies spaced repetition, retrieval, and learner autonomy as recurring affordances of digital flashcard systems (Teymouri, 2024). However, the 3R logic can be implemented entirely with paper cards, notebooks, slides, mini-whiteboards, or teacher questioning. This is important because the pedagogical mechanism is retrieval over time, not the application itself. A visually attractive platform that repeatedly shows answers without requiring recall does not automatically produce retrieval practice.

The same principle applies to gamified review. Points, timers, leaderboards, and competition can increase engagement, but they should not replace the cognitive requirements of the task. A game becomes relevant to durable vocabulary learning when learners must actually retrieve target information, receive corrective feedback, and meet the words again after a meaningful interval. The framework therefore allows teachers to use familiar tools while evaluating them according to learning mechanism rather than novelty.

4.7. Boundaries and cautions

Three cautions are necessary. First, retrieval difficulty must be calibrated. Retrieval that is too easy may add little desirable effort, whereas retrieval that is consistently impossible can produce frustration and little successful practice. Teachers should therefore provide an initial learning phase strong enough to make later retrieval possible, and they should shorten the interval or strengthen cues for repeatedly failed items. Second, spacing should not be interpreted as delaying all feedback. Students need access to the correct answer after a genuine attempt, particularly when the response is wrong. The purpose of spacing is to distribute learning opportunities, not to leave errors unresolved.

Third, retrieval-and-spacing routines cannot replace rich language use. Deliberate retrieval is well suited to stabilizing form-meaning links and selected lexical features, but vocabulary knowledge grows through multiple encounters in varied contexts. Research showing that retrieval practice can support meaningful learning beyond rote recall (Karpicke & Blunt, 2011) is encouraging, yet a school vocabulary program still needs reading, listening, speaking, writing, and attention to collocation and word family relationships. The 3R framework should therefore function as a retention layer added to communicative and text-based teaching rather than as a return to isolated memorization.

Finally, much of the controlled retrieval and spacing evidence comes from university learners outside Vietnam. The framework is therefore a theoretically grounded adaptation, not proof of effectiveness for Vietnamese adolescents. Variables such as proficiency, class size, test pressure, item difficulty, teacher workload, and students' prior learning strategies may

moderate outcomes. These factors should be treated as empirical questions in future classroom research.

5. CONCLUSION

This review argues that one of the most practical improvements to upper-secondary EFL vocabulary teaching is to plan for future retrieval at the moment a word is first taught. Vietnamese research already indicates the importance of vocabulary, the lexical demands of school materials, and the need for principled vocabulary pedagogy (N. T. M. Le & Dinh, 2022; Lien et al., 2024; D. H. Tran, 2026; D. V. Vu & Peters, 2021). Retrieval-practice research adds a complementary insight: durable knowledge is strengthened when learners are required to reconstruct previously learned information rather than repeatedly restudy it (Karpicke & Roediger, 2008; Nakata, 2017; Rowland, 2014). Spacing research further shows that these retrieval opportunities are more useful for long-term retention when they are distributed across time and revisited cumulatively (Kim & Webb, 2022, 2023; Nakata et al., 2021; Nakata et al., 2023).

The proposed 3R framework converts these findings into a classroom sequence. Retrieve requires an effortful attempt before the answer is shown. Revisit returns selected items across later lessons using cumulative and adjustable intervals. Recontextualize moves successful recall into new sentences, collocations, speaking, and writing so that vocabulary knowledge becomes usable rather than merely recognizable. The model is intentionally low-tech and low-burden; digital systems can support it, but the central mechanism can be implemented with ordinary classroom tools.

The framework also suggests a research agenda. Future Vietnamese studies could compare conventional end-of-unit review with 3R-based distributed retrieval, examine immediate and delayed outcomes separately, and test both receptive and productive vocabulary knowledge. Studies should report actual retrieval schedules, feedback procedures, item selection criteria, and time-on-task so that findings can be transferred across classrooms. A further priority is to investigate whether cumulative retrieval can support examination performance without narrowing vocabulary learning to recognition-based multiple-choice practice.

The main pedagogical conclusion is modest but consequential: vocabulary review should not be concentrated immediately after teaching or immediately before a test. By engineering short, successful, and increasingly contextualized retrieval opportunities across time, teachers can turn ordinary review into a deliberate system for durable lexical access.

REFERENCES

[1]. Huyen, D. H., & Nghi, T. T. (2026). Effects of Kahoot!-mediated vocabulary instruction on vocabulary gains and learner attitudes: A mixed-methods study in a Vietnamese public high school. *English Language Teaching Educational Journal, 9*(1), 173–185. https://doi.org/10.12928/eltej.v9i1.15279

[2]. Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*(6018), 772–775. https://doi.org/10.1126/science.1199327

[3]. Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966–968. https://doi.org/10.1126/science.1152408

[4]. Kim, S. K., & Webb, S. (2022). The effects of spaced practice on second language learning: A meta-analysis. *Language Learning, 72*(1), 269–319. https://doi.org/10.1111/lang.12479

[5]. Kim, S. K., & Webb, S. (2023). Does spaced practice have the same effects on different second language vocabulary learning activities? Fill-in-the-blanks versus flashcards. *The Modern Language Journal, 107*(4), 944–964. https://doi.org/10.1111/modl.12879

[6]. Le, N. T. M., & Dinh, H. T. (2022). Vocabulary coverage in a high school Vietnamese EFL textbook: A corpus-based preliminary investigation. *Vietnam Journal of Education, 6*(2), 102–113. https://doi.org/10.52296/vje.2022.187

[7]. Le, T. V., Mai, T. M., Tran, T. L. N., & Nguyen, L. T. (2024). Teacher perspectives about English language teaching reforms in Vietnamese high schools. *Ho Chi Minh City Open University Journal of Science - Social Sciences, 14*(4), 3–18. https://doi.org/10.46223/HCMCOUJS.soci.en.14.4.2781.2024

[8]. Lien, N. N., Mai, N. H., & Trang, N. H. (2024). Vocabulary in English textbooks for Vietnamese upper-secondary students: A comparative analysis of reading passages. *TESL-EJ, 28*(2). https://doi.org/10.55593/ej.28110a10

[9]. Nakata, T. (2017). Does repeated practice make perfect? The effects of within-session repeated retrieval on second language vocabulary learning. *Studies in Second Language Acquisition, 39*(4), 653–679. https://doi.org/10.1017/S0272263116000280

[10]. Nakata, T., Suzuki, Y., & He, X. (2023). Costs and benefits of spacing for second language vocabulary learning: Does relearning override the positive and negative effects of spacing? *Language Learning, 73*(3), 799–834. https://doi.org/10.1111/lang.12553

[11]. Nakata, T., Tada, S., McLean, S., & Kim, Y. A. (2021). Effects of distributed retrieval practice over a semester: Cumulative tests as a way to facilitate second language vocabulary learning. *TESOL Quarterly, 55*(1), 248–270. https://doi.org/10.1002/tesq.596

[12]. Nguyen, T. B. T. (2024). Task-based language learning in Vietnamese EFL high school classrooms: Teachers' roles in the during-task phase and their beliefs. *VNU Journal of Foreign Studies, 40*(6), 152–170. https://doi.org/10.63023/2525-2445/jfs.ulis.5271

[13]. Qu, K., Liu, T., Qiao, Y., & Wang, P. (2024). The facilitative effect of the keyword mnemonic on L2 vocabulary retrieval practice. *Heliyon, 10*(3), Article e25212. https://doi.org/10.1016/j.heliyon.2024.e25212

[14]. Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. https://doi.org/10.1037/a0037559

[15]. Teymouri, R. (2024). Recent developments in mobile-assisted vocabulary learning: A mini review of published studies focusing on digital flashcards. *Frontiers in Education, 9*, Article 1496578. https://doi.org/10.3389/feduc.2024.1496578

[16]. Tran, D. H. (2026). Students' attitudes towards English vocabulary learning at an upper secondary school in the Central Highlands of Vietnam. *Ho Chi Minh City University of Education Journal of Science, 23*(1), 92–102. https://doi.org/10.54607/hcmue.js.23.1.4957(2026)

[17]. Tran, D. T., & Le, H. T. (2023). The impact of Quizlet on vocabulary improvement: A case study in Binh Duong Province secondary schools. *World Journal of English Language, 13*(6), 235–242. https://doi.org/10.5430/wjel.v13n6p235

[18]. Tran, M. D., & Nguyen, T. M. L. (2022). Improving students' vocabulary retention through using Quizlet: An action research at high school. *TNU Journal of Science and Technology, 227*(4), 170–177. https://doi.org/10.34238/tnu-jst.5647

[19]. Uchihara, T. (2023). How does the test modality of weekly quizzes influence learning the spoken forms of second language vocabulary? *TESOL Quarterly, 57*(2), 595–617. https://doi.org/10.1002/tesq.3176

[20]. Vu, D. V., & Peters, E. (2021). Vocabulary in English language learning, teaching, and testing in Vietnam: A review. *Education Sciences, 11*(9), Article 563. https://doi.org/10.3390/educsci11090563

[21]. Vu, H. H. (2025). The shift of the English paper from compulsory to elective status in the national upper-secondary school exam in Vietnam: Impacts on teaching and assessment from teachers' perspectives. *VNU Journal of Foreign Studies, 41*(3), 1–13. https://doi.org/10.63023/2525-2445/jfs.ulis.5411