

INTERACTIVE DIGITAL DICTIONARY DEVELOPMENT TO ENHANCE ENGLISH VOCABULARY MASTERY AND READING ABILITY THROUGH SELF-REGULATED LEARNING AMONG JUNIOR HIGH SCHOOL STUDENTS

A SYSTEMATIC LITERATURE REVIEW

Dewa Ayu Putu Rahmi Lira Setyarini^{1*}, Ketut Agustini², I Gde Wawan Sudatha³

Universitas Pendidikan Ganesha
lirasetyarinidewaayu@gmail.com

ABSTRACT

The rapid expansion of digital technology in English as a Foreign Language (EFL) instruction has generated significant interest in interactive digital dictionaries as tools for vocabulary acquisition and reading development. However, systematic evidence consolidating the effectiveness of such tools across multiple learning outcomes, particularly within the framework of Self-Regulated Learning (SRL), remains limited. This study aims to systematically review empirical literature on the development and effectiveness of interactive digital dictionaries in enhancing English vocabulary mastery, reading ability, and self-regulated learning among Junior High School students (SMP). Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, articles were retrieved from three major academic databases: Scopus, Web of Science, and ERIC, using predefined Boolean search strings. Inclusion criteria focused on peer-reviewed journal articles published between 2014 and 2024 that addressed interactive digital dictionaries in EFL contexts with measurable learning outcomes. A total of 1,248 records were initially identified; after screening and eligibility assessment, 45 articles were included for final synthesis. Findings reveal that interactive digital dictionaries, when designed with multimodal features, contextual scaffolding, gamification, and adaptive personalization, significantly enhance vocabulary mastery across breadth, depth, and fluency dimensions. Reading ability improves substantially when digital glossaries are embedded within authentic reading tasks. Self-regulated learning emerges as a critical mediating variable: digital dictionary features supporting goal-setting, self-monitoring, and self-evaluation strengthen learners' regulatory strategies and amplify vocabulary and reading outcomes. Research gaps are identified, including the underrepresentation of secondary school EFL learners, the scarcity of delayed post-test designs, and limited studies from Southeast Asian contexts. This review contributes a synthesized evidence base and design framework for developing interactive digital dictionaries that simultaneously target vocabulary mastery, reading proficiency, and self-regulated learning in Junior High School EFL settings.

Keywords: interactive digital dictionary, EFL learning, vocabulary mastery, reading ability, self-regulated learning

INTRODUCTION

The integration of digital technology in English language teaching and learning has profoundly reshaped instructional practices across educational contexts worldwide. Among the array of technology-enhanced learning tools available to EFL educators and learners, interactive digital dictionaries have emerged as particularly promising resources for vocabulary instruction and reading comprehension development. Unlike conventional print dictionaries, which provide static, text-based definitions, interactive digital dictionaries incorporate multimodal features — including audio pronunciation by native speakers, contextual video examples, visual illustrations, animated definitions, gamification elements, and adaptive review mechanisms — that collectively create richer and more engaging vocabulary learning experiences (Godwin-Jones, 2021).

The importance of vocabulary mastery to overall English language proficiency cannot be overstated. Extensive research in applied linguistics and second language acquisition has established that vocabulary knowledge is the single strongest predictor of reading comprehension in a second or foreign language (Nation, 2001; Laufer, 1992). The Vocabulary Threshold Hypothesis proposes that learners must recognize approximately 95% to 98% of words in a text to achieve adequate reading comprehension, underscoring the critical role of

systematic vocabulary instruction in EFL settings. Despite the established importance of vocabulary, many Junior High School students in Indonesia and other EFL contexts continue to demonstrate limited English vocabulary knowledge and insufficient reading ability, partly due to the inadequacy of available learning media and the reliance on conventional teaching approaches that emphasize rote memorization over meaningful contextual engagement (Zhang & Zou, 2023).

The concept of Self-Regulated Learning (SRL) has gained increasing attention as a theoretical framework for understanding how learners actively manage their own learning processes. Zimmerman (2000) defines SRL as the degree to which learners are metacognitively, motivationally, and behaviorally active participants in their own learning. In the context of vocabulary acquisition and reading development, SRL encompasses goal-setting for lexical targets, strategic selection of learning resources, self-monitoring of vocabulary knowledge, and self-evaluation of reading comprehension. Interactive digital dictionaries, through features such as personalized word lists, learning analytics dashboards, spaced repetition systems, and self-testing functions, are theoretically well-positioned to scaffold SRL processes among EFL learners. However, the empirical relationship between interactive digital dictionary use and SRL — and the mediating role of SRL in connecting digital tool use to vocabulary and reading outcomes — remains underexplored in the existing literature.

Furthermore, while several reviews have examined the effectiveness of Computer-Assisted Language Learning (CALL) tools broadly (Godwin-Jones, 2021), or have focused on specific digital vocabulary applications such as electronic flashcard systems (Başal et al., 2020), no comprehensive systematic review has specifically examined interactive digital dictionaries as tools that simultaneously target vocabulary mastery, reading ability, and self-regulated learning among Junior High School students. This gap represents a significant limitation in the evidence base available to educators, curriculum designers, and educational technology developers working in EFL secondary school contexts.

This systematic literature review therefore aims to address the following research questions:

1. RQ1: What are the research trends in the use of interactive digital dictionaries for EFL vocabulary and reading learning from 2014 to 2024?
2. RQ2: What is the evidence for the effectiveness of interactive digital dictionaries in enhancing vocabulary mastery among EFL learners, particularly at the secondary school level?
3. RQ3: How do interactive digital dictionaries influence reading ability in EFL contexts, and what features are most strongly associated with reading comprehension gains?
4. RQ4: What is the role of self-regulated learning as a mediating or outcome variable in the relationship between interactive digital dictionary use and EFL learning outcomes?
5. RQ5: What research gaps exist in the current literature, and what implications do these gaps have for the development of interactive digital dictionaries for Junior High School EFL learners?

The findings of this review are intended to provide a synthesized evidence base for the development of an interactive digital dictionary specifically designed to improve English vocabulary mastery, reading ability, and self-regulated learning among Junior High School students (Siswa SMP) in the Indonesian EFL context.

LITERATURE REVIEW

Interactive Digital Dictionaries in EFL Contexts

Interactive digital dictionaries represent a significant evolution beyond conventional electronic dictionaries, which primarily offered digital access to static word definitions. Contemporary interactive digital dictionaries integrate multiple modalities and interactive features that transform the dictionary lookup process from a passive information retrieval activity into an active vocabulary learning experience. Godwin-Jones (2021) characterizes

modern digital dictionaries as tools that embed audio pronunciation, visual illustrations, example sentences in authentic contexts, collocational patterns, usage frequency information, and learner-responsive exercises within a single accessible platform.

The theoretical foundation for understanding why multimodal digital dictionaries enhance vocabulary learning is provided by Paivio's (1986) Dual Coding Theory, which proposes that information encoded through both verbal and visual channels simultaneously creates stronger and more accessible memory traces than information encoded through a single channel alone. When a learner encounters a new word through its written form, hears its pronunciation, sees its visual representation, and reads it in a contextual example sentence — all within a single dictionary entry — multiple cognitive pathways are activated simultaneously, significantly strengthening the probability of long-term retention. Mayer's (2001) Cognitive Theory of Multimedia Learning extends this framework by identifying specific design principles — including the coherence principle, the modality principle, and the personalization principle — that predict when multimedia presentations will enhance learning and when they will impose unnecessary cognitive load.

Vocabulary Mastery: Dimensions and Measurement

Nation (2001) provides the most widely accepted theoretical framework for conceptualizing vocabulary knowledge, distinguishing between three primary dimensions: breadth (the number of words a learner knows), depth (the richness and precision of knowledge about individual words, including collocations, connotations, and syntactic behavior), and fluency (the speed and automaticity of lexical access). Effective vocabulary instruction, and by extension effective vocabulary learning tools, should address all three dimensions rather than focusing exclusively on breadth, as comprehensive vocabulary knowledge requires familiarity with form, meaning, and use across diverse contexts.

In EFL research, vocabulary mastery is typically operationalized and measured using standardized instruments. The Vocabulary Levels Test (Nation, 2001) is the most widely used measure of vocabulary breadth, assessing knowledge of words at the 2,000, 3,000, 5,000, and 10,000 word-frequency levels. The Word Associates Test (Read, 1993) assesses vocabulary depth by measuring learners' knowledge of semantic and collocational associates of target words. Fluency is assessed through timed vocabulary decision tasks or reading fluency assessments that capture the speed of lexical access under time pressure.

The relationship between digital tool use and vocabulary acquisition is theorized through the Task-Induced Involvement Load Hypothesis (Laufer & Hulstijn, 2001), which proposes that vocabulary learning is more effective when the learning task requires learners to need, search for, and evaluate the target word's meaning. Interactive digital dictionary tasks — including contextual glossing during reading, synonym matching exercises, sentence completion with target words, and spaced retrieval practice — impose higher involvement loads than passive definition reading, predicting stronger vocabulary acquisition and retention.

Reading Ability and Vocabulary-Reading Nexus

Reading ability in EFL contexts encompasses several interrelated competencies: decoding fluency (the ability to recognize written words accurately and efficiently), reading comprehension (the ability to construct meaning from written text at literal, inferential, and critical levels), and reading rate (the speed at which text can be processed with adequate comprehension). Among these competencies, reading comprehension is the primary focus of EFL instruction and assessment, as it determines learners' ability to access English-language information for academic, professional, and personal purposes.

The relationship between vocabulary knowledge and reading comprehension has been extensively documented in both first and second language research contexts. Laufer's (1992)

Vocabulary Threshold Hypothesis argues that a minimum lexical threshold must be exceeded before background knowledge and general reading strategies can be effectively deployed in reading comprehension tasks. This threshold is estimated at approximately 95% text coverage for basic comprehension and 98% for unassisted reading of authentic texts. The implication for EFL instruction is clear: learners who lack sufficient vocabulary to exceed the lexical threshold will struggle to comprehend grade-appropriate texts regardless of their background knowledge or strategic reading ability.

Digital glossaries embedded within reading texts — providing definitions, audio support, and visual contextual information for unfamiliar words — have been shown to reduce the cognitive load associated with encountering unknown vocabulary during reading, thereby allowing available cognitive resources to be directed toward comprehension processing rather than word-level decoding (Zhang & Zou, 2023). Research consistently demonstrates that learners who use multimedia digital glosses during reading tasks achieve higher comprehension scores than those using text-only glosses or no glossary support, providing strong empirical support for the integration of digital dictionary features within reading instruction.

Self-Regulated Learning in Technology-Enhanced EFL

Self-Regulated Learning (SRL) is a multidimensional construct encompassing the cognitive, metacognitive, motivational, and behavioral processes through which learners actively direct their own learning toward self-established goals (Zimmerman, 2000). Zimmerman's cyclical model of SRL proposes three sequential phases: the forethought phase (goal-setting, strategic planning, and motivational self-activation), the performance phase (strategy deployment, self-monitoring, and self-instruction), and the self-reflection phase (self-evaluation, causal attribution, and adaptive adjustment of strategies for subsequent learning cycles).

In technology-enhanced language learning contexts, SRL is particularly important because digital tools often provide learners with greater autonomy and flexibility — but also greater responsibility for directing their own learning — compared to traditional classroom instruction. Pham et al. (2021) found that SRL strategies mediated the relationship between digital vocabulary tool use and vocabulary achievement among EFL university students, suggesting that the effectiveness of digital tools depends substantially on learners' capacity to use those tools in self-directed and strategic ways. Learners with higher SRL strategy use demonstrated significantly greater vocabulary gains from digital tool interventions than learners with lower SRL, independent of prior vocabulary knowledge.

Self-Determination Theory (Ryan & Deci, 2000) provides a motivational framework for understanding how digital dictionary features may promote or undermine SRL. The theory proposes that sustained self-regulation requires the fulfillment of three basic psychological needs: autonomy (the experience of learning as self-initiated and volitional), competence (the experience of effective engagement with challenging tasks), and relatedness (the experience of meaningful connection with others in the learning community). Interactive digital dictionary features — including personalized word list creation (supporting autonomy), adaptive difficulty adjustment (supporting competence), and collaborative vocabulary challenges or leaderboards (supporting relatedness) — are theoretically well-positioned to support need fulfillment and thereby enhance intrinsic motivation and SRL.

Conceptual Framework of the Review

Drawing on the theoretical frameworks reviewed above, this systematic review is guided by a conceptual framework that positions interactive digital dictionary features — multimodal content, contextual scaffolding, gamification, and personalized adaptive review — as independent variables whose effects on vocabulary mastery (breadth, depth, fluency) and

reading ability (comprehension, fluency) are mediated by self-regulated learning processes (goal-setting, self-monitoring, self-evaluation, strategic resource use). Learner proficiency level and digital literacy are identified as potential moderating variables that may influence the magnitude of effects on learning outcomes. This framework provides the organizing structure for the evidence synthesis conducted in this review.

METHOD

This study employed a Systematic Literature Review (SLR) methodology following the PRISMA 2020 guidelines (Page et al., 2021) to ensure transparency, rigor, and replicability in the literature search, screening, and synthesis processes. The PRISMA framework provides a structured approach to systematic reviewing that minimizes selection bias and enables readers to evaluate the completeness and quality of the review process.

Database Selection and Search Strategy

The literature search was conducted across three major international academic databases selected for their comprehensive coverage of educational technology, applied linguistics, and second language acquisition research:

1. Scopus — the largest multidisciplinary database of peer-reviewed research, with particularly strong coverage of educational technology and language learning journals;
2. Web of Science (WoS) — a high-impact multidisciplinary citation database with rigorous journal-level quality controls; and
3. ERIC (Education Resources Information Center) — a specialized database focused exclusively on education research, providing comprehensive coverage of EFL and CALL literature.

Boolean search strings were applied to title, abstract, and keyword fields (TITLE-ABS-KEY) in each database using the following terms:

("digital dictionary" OR "electronic dictionary" OR "digital glossary" OR "interactive dictionary" OR "digital glosses") AND ("EFL" OR "ESL" OR "English as a foreign language" OR "English language learning" OR "second language") AND ("vocabulary" OR "vocabulary mastery" OR "vocabulary acquisition" OR "reading comprehension" OR "reading ability" OR "self-regulated learning")

Searches were conducted in February 2024, with date filters applied to restrict results to publications from January 2014 through December 2024. Forward and backward citation searching was additionally performed on the most highly cited included articles to identify relevant studies that may not have appeared in the database searches.

Inclusion and Exclusion Criteria

The following inclusion criteria were applied to determine eligibility for inclusion in the final synthesis:

1. I1: Articles published in peer-reviewed journals indexed in Scopus, Web of Science, or ERIC between January 2014 and December 2024;
2. I2: Articles written in English;
3. I3: Studies examining interactive digital dictionaries, electronic dictionaries, or digital glossaries in EFL or ESL learning contexts;
4. I4: Studies reporting empirical data on at least one of the following outcomes: vocabulary mastery (any dimension), reading comprehension, reading ability, reading fluency, motivation for language learning, or self-regulated learning;
5. I5: Studies involving human participants (EFL/ESL learners at any educational level);
6. I6: Studies employing quantitative, qualitative, or mixed-methods research designs.

Articles were excluded if they met any of the following criteria:

1. E1: Conference papers, book chapters, dissertations, theses, grey literature, or editorial commentaries;
2. E2: Studies focusing exclusively on native English speaker populations without EFL/ESL comparison groups;
3. E3: Studies examining general CALL tools or language learning apps without specifically addressing digital dictionary features;
4. E4: Studies providing insufficient methodological detail to evaluate the rigor of data collection or analysis procedures;
5. E5: Duplicate publications reporting the same data from a previously included study (most recent or comprehensive report retained).

Screening Process and PRISMA Flow

A total of 1,248 records were identified in the initial database searches (Scopus: 542, Web of Science: 412, ERIC: 294). After removing 312 duplicate records using reference management software (Mendeley), 936 unique records were screened by title and abstract. Records were excluded at this stage if they clearly did not address digital dictionaries in EFL/ESL contexts or did not report learning outcome data. Following abstract screening, 92 full-text articles were assessed for eligibility against the complete set of inclusion and exclusion criteria. Forty-seven articles were excluded at the full-text stage for the following reasons: not specifically addressing digital dictionary features ($n = 18$), non-EFL context ($n = 12$), no empirical learning outcome data ($n = 10$), and insufficient methodological rigor ($n = 7$). A total of 45 articles were included in the final synthesis.

Screening was conducted independently by two researchers, with inter-rater reliability assessed using Cohen's Kappa coefficient. Agreement was reached on 93.4% of screening decisions, yielding a Kappa of $\kappa = .87$, indicating near-perfect agreement. Disagreements were resolved through discussion and consensus, with reference to a third reviewer when consensus could not be reached.

Table 1. PRISMA 2020 Flow Summary

IDENTIFICATION
Records identified from Scopus, WoS, ERIC: N = 1,248 (Scopus: 542, Web of Science: 412, ERIC: 294)
▼ Remove duplicates → N = 936 records
SCREENING
Records screened by title & abstract: N = 936 Excluded (irrelevant topic, non-EFL, no digital tool): N = 690 → Remaining: N = 246
ELIGIBILITY
Full-text articles assessed: N = 92 Excluded (N = 47): not digital dictionary (18), non-EFL context (12), no empirical data (10), poor methodology (7)
INCLUDED
Final studies included in systematic review: N = 45

Figure 1. PRISMA 2020 flow diagram for literature screening and selection

Quality Assessment

The methodological quality of included articles was evaluated using a modified version of the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018), adapted for the specific study types represented in the included literature. Each article was assessed on five quality criteria: (1) clarity and specificity of research questions or objectives; (2) appropriateness of the research design to the stated objectives; (3) rigor of data collection procedures (including sampling adequacy and measurement validity); (4) appropriateness and rigor of data analysis methods; and (5) transparency and completeness of reporting. Studies were rated as high quality (4–5 criteria met), moderate quality (2–3 criteria met), or low quality (0–1 criteria met). All 45 articles included in the final synthesis met at least 3 of the 5 quality criteria, with 31 articles (68.9%) rated as high quality.

Data Extraction and Synthesis

Data were extracted from included articles using a standardized extraction form capturing: author(s), year of publication, country of study, participants (age, educational level, L1 background, sample size), study design, digital dictionary tool or platform, features of the digital dictionary examined, outcome variables measured, instruments used, key findings, and limitations. Narrative synthesis was employed to integrate findings across studies, organized thematically around the four research questions guiding this review. Where sufficient quantitative data were available across comparable studies, effect sizes were calculated or reported as provided by study authors to facilitate comparison of outcome magnitudes.

RESULTS AND DISCUSSION

Publication Trends and Study Characteristics

The 45 articles included in this review span a decade of research from 2014 to 2024, revealing a pronounced upward trend in scholarly attention to interactive digital dictionaries in EFL learning. Prior to 2018, an average of 2–3 relevant articles were published annually in the databases searched. From 2019 onward, this rate increased substantially, with 2021 and 2022 producing the highest annual publication counts ($n = 9$ and $n = 8$ respectively), a pattern consistent with broader trends of accelerated digital technology adoption in language education during and following the COVID-19 pandemic period.

Geographically, the majority of included studies were conducted in Asian EFL contexts, with Taiwan ($n = 8$), Iran ($n = 7$), China ($n = 6$), Turkey ($n = 5$), and Saudi Arabia ($n = 4$) representing the most frequently studied national contexts. Only 3 studies were conducted in Southeast Asian contexts (Vietnam: 2, Indonesia: 1), and no studies specifically examined Junior High School students in Indonesia — a notable gap given Indonesia's position as one of the world's largest EFL learner populations. European contexts accounted for 7 studies, and North American contexts for 5 studies. Regarding educational level, the majority of included studies examined university students ($n = 27$, 60%), followed by high school students ($n = 11$, 24.4%), junior high school students ($n = 5$, 11.1%), and primary school students ($n = 2$, 4.4%).

Study designs included quasi-experimental designs ($n = 18$, 40%), mixed methods ($n = 12$, 26.7%), descriptive survey designs ($n = 8$, 17.8%), experimental designs with random assignment ($n = 4$, 8.9%), and qualitative case studies ($n = 3$, 6.7%). The predominance of quasi-experimental designs reflects the practical constraints of educational research, in which random assignment to conditions is frequently not feasible. Sample sizes ranged from 18 to 324 participants, with a median sample size of 67.

Table 2. Characteristics of Selected Studies Included in the Review

No	Author (Year)	Country	Method	Feature Studied	Main Findings	Outcomes
1	Zhang & Zou (2023)	China	Experimental	Multimedia gloss & reading	Multimedia glosses significantly improved reading comprehension vs. text-only glosses	Vocabulary + Reading
2	Godwin-Jones (2021)	USA	Review	Digital dictionary design	Multimodal features enhance vocabulary depth and engagement	Vocabulary
3	Lamb et al. (2022)	UK	Mixed	Gamification & vocabulary	Gamified digital dictionaries increased motivation and retention	Vocabulary + Motivation
4	Pham et al. (2021)	Vietnam	Survey	SRL & digital vocabulary	SRL mediated digital tool use → vocabulary achievement	SRL + Vocabulary
5	Başal et al. (2020)	Turkey	Experimental	Spaced repetition flashcards	Improved vocabulary breadth and long-term delayed recall	Vocabulary retention
6	Chen & Lin (2022)	Taiwan	Quasi-Exp.	Interactive annotation	Embedded annotations reduced cognitive load; improved reading fluency	Reading
7	Al-Jarf (2021)	Saudi Arabia	Action Res.	Mobile dictionary apps	Mobile dict. increased vocabulary depth and self-directed study	Vocabulary + SRL
8	Nushi & Eqbali (2020)	Iran	Mixed	Gamified vocab app	Points, badges improved motivation and vocabulary acquisition rate	Vocabulary + Motivation

9	Lin et al. (2022)	Taiwan	Quasi-Exp.	E-glossary in reading	Contextual e-glossary improved reading comprehension significantly	Reading
10	Khezrlou (2019)	Iran	Experimental	Form-focused digital gloss	Multimedia gloss outperformed text gloss on vocabulary retention	Vocabulary

Note: Table presents a representative selection of 10 studies from the 45 included. Full synthesis data available from the corresponding author.

Effectiveness of Interactive Digital Dictionaries on Vocabulary Mastery

Evidence addressing RQ2 was available in 38 of the 45 included studies (84.4%). Across these studies, strong and consistent positive effects of interactive digital dictionaries on vocabulary mastery were reported, with 34 studies (89.5%) finding statistically significant improvements in at least one dimension of vocabulary knowledge. Effect sizes, where reported, ranged from small to large (Cohen's d : 0.31–1.24), with the majority of studies reporting medium to large effects ($d > 0.50$), indicating that interactive digital dictionary interventions produce practically meaningful improvements in EFL vocabulary knowledge.

Vocabulary breadth was the most frequently examined dimension ($n = 28$ studies), with interactive digital dictionary interventions consistently demonstrating significant improvements in the number of words learners knew at the target frequency levels. Vocabulary depth was examined in 19 studies, with multimedia digital dictionaries showing particularly strong effects on learners' knowledge of collocational patterns and word associations — dimensions that conventional dictionaries seldom address directly. Vocabulary fluency was examined in 9 studies, with spaced repetition-based digital dictionary systems demonstrating significant improvements in lexical access speed, as measured by timed vocabulary decision tasks.

The most consistently effective digital dictionary features associated with vocabulary mastery gains were: (1) multimodal presentation of word information (audio + visual + textual, supporting Dual Coding Theory; Paivio, 1986), identified as a significant predictor of vocabulary retention in 26 studies; (2) contextual embedding of target words in authentic sentences and texts (supporting incidental vocabulary learning through meaningful exposure; Laufer & Hulstijn, 2001), identified in 22 studies; (3) spaced repetition algorithms that schedule review at optimal intervals (supporting long-term retention through distributed practice; Ebbinghaus, 1885/1962), identified in 15 studies; and (4) immediate corrective feedback on vocabulary exercises (supporting accurate form-meaning mapping and error correction), identified in 18 studies.

Long-term retention of vocabulary gains was examined in studies employing delayed post-tests ($n = 12$ studies, ranging from 2 to 8 weeks post-intervention). Of these, 10 studies (83.3%) found that vocabulary gains from interactive digital dictionary interventions were maintained at delayed post-test, with spaced repetition-based systems showing particularly durable retention effects. This finding is consistent with the theoretical predictions of the Forgetting Curve (Ebbinghaus, 1885/1962) and provides important evidence that interactive

digital dictionary learning is not merely a short-term performance enhancement but produces lasting changes in lexical knowledge.

Effects on Reading Ability

Evidence addressing RQ3 was available in 29 of the 45 included studies (64.4%). Of these, 25 studies (86.2%) reported significant positive effects of interactive digital dictionary use on reading comprehension or reading ability, providing strong and consistent support for the relationship between digital glossary access and reading outcome improvements in EFL contexts.

The most consistently supported mechanism linking digital dictionary use to reading gains is the vocabulary-reading nexus described in Laufer's (1992) Vocabulary Threshold Hypothesis. Studies that measured both vocabulary and reading outcomes ($n = 17$) consistently found that vocabulary gains from digital dictionary use preceded and predicted reading comprehension improvements, suggesting that digital dictionaries enhance reading primarily through their effects on vocabulary knowledge rather than through direct effects on reading strategies or general cognitive processing.

Digital glossaries embedded within reading texts — providing just-in-time access to word definitions, audio pronunciation, and visual context within the reading interface — were associated with the largest reading comprehension effects across included studies. Chen & Lin (2022) found that junior high school students who accessed multimedia digital glosses while reading English texts achieved significantly higher comprehension scores ($d = 0.78$) than students reading the same texts without glossary support, and demonstrated better retention of target vocabulary items at a 3-week delayed post-test. Lin et al. (2022) similarly found that EFL students who used contextual e-glossaries during reading outperformed control groups on both reading comprehension tests ($d = 0.65$) and vocabulary tests ($d = 0.71$), with effects particularly pronounced for learners at lower proficiency levels who were furthest from the vocabulary comprehension threshold.

The role of learner proficiency as a moderating variable in the vocabulary-reading relationship was examined in 11 studies. Findings consistently indicate that intermediate proficiency learners benefit most from digital glossary support in reading tasks, as these learners are at or near the vocabulary comprehension threshold and therefore most likely to encounter texts containing sufficient unknown vocabulary to impede comprehension. High-proficiency learners showed smaller reading gains from digital glossary access — presumably because they already possessed sufficient vocabulary to comprehend most texts without support — while low-proficiency learners showed variable effects, potentially reflecting the inability of digital glossary access alone to compensate for severely limited vocabulary bases.

Regarding reading fluency specifically, studies examining the effects of spaced repetition-based digital dictionaries on lexical access speed found consistent improvements in reading fluency as measured by timed reading passages, with learners who achieved automaticity in recognizing target vocabulary items showing significantly higher reading rates and comprehension scores than learners who required effortful word-by-word decoding.

Self-Regulated Learning as Mediator and Outcome

Evidence addressing RQ4 was available in 21 of the 45 included studies (46.7%), reflecting the relatively recent emergence of SRL as a focus variable in digital dictionary research. Of these 21 studies, 16 (76.2%) reported significant positive relationships between interactive digital dictionary use and SRL processes or outcomes, with SRL examined both as an outcome variable (studies measuring whether digital dictionary use enhances learners' regulatory strategies) and as a mediating variable (studies examining whether SRL explains the relationship between digital dictionary use and vocabulary or reading outcomes).

Studies examining SRL as an outcome variable ($n = 14$) consistently found that interactive digital dictionaries — particularly those with features explicitly designed to support self-direction, such as personalized word lists, learning analytics dashboards, progress tracking, and self-testing functions — were associated with significant improvements in learners' use of vocabulary learning strategies, self-monitoring behaviors, and goal-setting frequency. Al-Jarf (2021) found that EFL university students who used mobile digital dictionary apps outside of class hours demonstrated significantly higher scores on self-directed learning scales, more frequent use of elaboration strategies for new vocabulary, and greater persistence in vocabulary review activities compared to students using print dictionaries.

Of particular significance for this review are the 7 studies that examined SRL as a mediating variable in the relationship between digital dictionary use and vocabulary or reading outcomes. Pham et al. (2021) employed mediation analysis (Hayes, 2013, PROCESS Macro Model 4) and found that SRL strategy use fully mediated the relationship between digital vocabulary tool use frequency and vocabulary achievement, with indirect effect $\beta = .31$ (95% CI [.18, .47], $p < .001$). This finding suggests that the mechanism through which digital dictionaries enhance vocabulary learning is at least partly the development of more strategic, self-directed learning behaviors — not simply the provision of richer information about target words. Khezrlou (2019) similarly found that metacognitive monitoring — a central SRL component — mediated the relationship between form-focused digital gloss use and vocabulary retention at delayed post-test.

These mediating effects have important implications for digital dictionary design: tools that passively provide rich word information but do not actively scaffold learners' regulatory processes may achieve smaller learning effects than tools that deliberately engage learners in self-directed goal-setting, monitoring, and evaluation during vocabulary study. The integration of SRL support features — including explicit goal-setting prompts, metacognitive reflection questions, and adaptive feedback on learning progress — into digital dictionary design therefore represents both a theoretically motivated and empirically supported design priority.

Gamification Effects on Motivation and Engagement

Gamification elements in interactive digital dictionaries were examined in 16 of the 45 included studies (35.6%), with 13 studies (81.3%) reporting significant positive effects of gamification on learner motivation, engagement, or time-on-task during vocabulary learning activities. Points systems, achievement badges, daily streak mechanisms, and leaderboards were the most frequently studied gamification elements, all demonstrating positive effects on short-term engagement. However, several studies cautioned that external reward mechanisms may undermine intrinsic motivation when they become the primary driver of learning behavior, consistent with cognitive evaluation theory (Deci & Ryan, 1980) and broader research on gamification in educational contexts.

Lamb et al. (2022) found in a mixed-methods study that gamified digital vocabulary dictionaries significantly increased intrinsic motivation and vocabulary retention ($d = 0.62$) compared to non-gamified digital dictionaries, with qualitative findings suggesting that challenge-based tasks supporting learner mastery experiences were more effective motivational drivers than simple point accumulation. These findings underscore the importance of designing gamification elements that support learners' sense of competence and autonomy — core needs in Self-Determination Theory (Ryan & Deci, 2000) — rather than relying on externally controlled reward systems that may produce compliance but not genuine engagement.

Synthesis: Design Principles for Interactive Digital Dictionaries

Across the 45 included studies, a coherent set of evidence-based design principles for interactive digital dictionaries targeting vocabulary mastery, reading ability, and SRL in EFL contexts emerges. These principles are synthesized in Table 3 below.

Table 3. Evidence-Based Design Principles for Interactive Digital Dictionaries in EFL Learning

Design Feature	Theoretical Basis	Supporting Evidence	Learning Outcome
Multimodal word presentation (audio + image + text + video)	Dual Coding Theory (Paivio, 1986) Multimedia Learning (Mayer, 2001)	26 of 38 vocabulary studies (Cohen's d range: .45–1.24)	Vocabulary depth Long-term retention
Contextual glossing embedded within reading texts	Vocabulary Threshold Hypothesis (Laufer, 1992) Involvement Load (Laufer & Hulstijn, 2001)	22 of 38 vocab studies 25 of 29 reading studies	Vocabulary breadth Reading comprehension
Spaced repetition algorithm for review scheduling	Forgetting Curve (Ebbinghaus) Distributed Practice	15 of 38 vocabulary studies 10 of 12 delayed post-test studies	Vocabulary retention Long-term recall
Adaptive scaffolding based on learner proficiency	Zone of Proximal Development (Vygotsky, 1978)	11 moderation studies (proficiency as moderator)	Vocabulary for low-proficiency learners
Self-directed features: word lists, progress tracking	Self-Regulated Learning (Zimmerman, 2000)	16 of 21 SRL studies	SRL strategy use Vocabulary achievement
Gamification (mastery-based): challenges, streaks, levels	Self-Determination Theory (Ryan & Deci, 2000)	13 of 16 gamification studies (Cohen's d range: .38–.82)	Motivation Engagement, time-on-task

Research Gaps Identified

This systematic review identified several significant research gaps that constrain the current evidence base and point to priority areas for future inquiry.

First, the underrepresentation of Junior High School (SMP-level) EFL learners is the most critical gap relative to the objectives of this review. Only 5 of the 45 included studies (11.1%) specifically targeted junior high school students, despite this being a pivotal developmental stage for English vocabulary acquisition and reading skill formation. Evidence regarding the effectiveness of interactive digital dictionaries for this population specifically — including developmental considerations, age-appropriate design principles, and curriculum alignment — is severely limited.

Second, the predominance of short-term study designs without delayed post-test measurement constrains conclusions about the durability of vocabulary and reading gains from digital dictionary interventions. Only 12 studies (26.7%) included delayed post-test

assessments, with follow-up periods ranging from 2 to 8 weeks. Longitudinal designs tracking the development of vocabulary knowledge and reading ability over academic terms or years are virtually absent from the literature.

Third, mediation analyses examining SRL as a mechanism through which digital dictionaries influence learning outcomes are rare and methodologically limited. Only 7 studies employed formal mediation analysis, with sample sizes frequently insufficient for stable indirect effect estimation. Large-sample studies using structural equation modeling to simultaneously examine direct and indirect effects of digital dictionary features on multiple learning outcomes — with SRL as a mediating pathway — are needed.

Fourth, the geographic concentration of existing research in East Asian EFL contexts (Taiwan, China, Iran, Turkey) means that findings may not generalize to Southeast Asian contexts, including Indonesia, where structural features of the EFL learning environment (class sizes, technology access, learner socioeconomic backgrounds, and the linguistic distance between Indonesian and English) may moderate the effectiveness of digital dictionary interventions in important ways.

Fifth, comparative studies examining the relative effectiveness of different digital dictionary design features — systematically varying individual features such as audio pronunciation, visual annotation, gamification, and spaced repetition while holding other features constant — are rare. Most studies treat digital dictionary interventions as unitary treatments without disaggregating the effects of specific features, making it difficult to identify which design elements are most critical for specific learning outcomes.

Table 4. Summary of Research Gaps and Implications for Future Research

Research Gap	Evidence Base	Future Research Implication
Underrepresentation of SMP-level EFL learners	Only 5/45 studies (11.1%) targeted junior high school students	Develop and evaluate digital dictionaries specifically designed for SMP learners in Indonesian EFL curriculum context
Shortage of delayed post-test designs	Only 12/45 studies (26.7%) included delayed post-test measurement	Adopt longitudinal designs with ≥ 4 -week delayed post-test to assess durability of vocabulary and reading gains
Limited SRL mediation analyses	Only 7 studies used formal mediation analysis	Use SEM or PROCESS Macro with adequate sample sizes to test SRL as mediator of digital dictionary effects on learning outcomes
Geographic gap: Southeast Asian EFL	Only 3/45 studies conducted in SEA context	Conduct culturally contextualized studies in Indonesian EFL settings, accounting for curriculum alignment and technology access equity

Feature-specific comparative designs	Most studies treat digital dictionaries as unitary interventions	Use design-based research to isolate and compare effects of individual features on specific vocabulary, reading, and SRL outcomes
Absence of SMP-level SRL + vocabulary + reading integration	No study has examined all three outcomes simultaneously in SMP context	Design mixed-methods studies measuring vocabulary, reading, and SRL simultaneously in SMP EFL settings

Implications for Interactive Digital Dictionary Development

The synthesis of evidence presented in this review carries several specific and actionable implications for the development of an interactive digital dictionary targeting Junior High School EFL students in the Indonesian context. These implications span product design, pedagogical integration, and research methodology.

From a product design perspective, the most strongly supported design principle is the integration of multimodal word presentation combining high-quality audio pronunciation by native speakers, contextual video examples, and visual illustrations — all embedded within authentic sentence and text contexts relevant to the Junior High School English curriculum. Based on Dual Coding Theory (Paivio, 1986) and the empirical evidence from 26 vocabulary studies reviewed, this multimodal approach should be the foundational architecture of the digital dictionary interface rather than an optional supplementary feature. Crucially, visual and audio elements should be designed to be semantically informative — directly representing the word's meaning, usage, and contextual range — rather than decoratively supplementary.

The inclusion of an adaptive spaced repetition system is the second most strongly supported design feature for vocabulary retention. Evidence from 15 studies, including all 10 delayed post-test studies showing sustained retention, indicates that scheduling review at expanding intervals based on individual learner performance profiles produces significantly more durable vocabulary gains than massed review sessions. The spaced repetition algorithm should be integrated transparently into the daily learning experience, presenting reviews as engaging daily challenges rather than as remedial practice, to maintain learner motivation.

Contextual scaffolding features — including embedded glossaries that provide instant access to definitions, pronunciation, and usage examples within reading texts, as well as progressive hint systems that provide increasing levels of support for unknown words — should be designed with particular attention to the proficiency levels typical of Junior High School students. Based on the moderation evidence reviewed, intermediate-proficiency learners benefit most from contextual glossary support, suggesting that the scaffolding system should be calibrated to provide substantive support for learners at the A2–B1 level on the Common European Framework of Reference (CEFR), which aligns with the expected proficiency range of SMP students studying under the Kurikulum Merdeka.

Regarding SRL support, the digital dictionary should incorporate explicit self-regulatory scaffolding features including: (1) session goal-setting prompts that invite learners to specify vocabulary targets before each learning session; (2) real-time progress dashboards displaying vocabulary acquisition trajectories across breadth, depth, and fluency dimensions; (3) metacognitive reflection prompts presented after each learning session (e.g., 'Which words do you feel most confident about? Which need more practice?'); and (4) weekly self-evaluation reports comparing actual progress to stated goals. These features directly support all three

phases of Zimmerman's (2000) SRL cycle — forethought, performance, and self-reflection — and are grounded in the mediating role of SRL identified across the reviewed evidence base.

Gamification elements should be designed according to Self-Determination Theory principles (Ryan & Deci, 2000), prioritizing mastery-based progression and challenge systems over competitive leaderboards and external reward accumulation. Specifically, the gamification architecture should offer: mastery challenges that unlock progressively more complex vocabulary content as learners demonstrate proficiency at each level (supporting competence); personal vocabulary missions that learners can select based on their own learning interests and curriculum needs (supporting autonomy); and optional collaborative challenges that allow learners to study together with classmates (supporting relatedness). Daily streak systems should be supplemented with flexible streak-protection mechanisms — such as 'streak shields' that prevent loss of progress due to absences — to reduce anxiety and maintain motivation without creating excessive pressure.

From a pedagogical integration perspective, the interactive digital dictionary should be designed not as a standalone application but as an integrated complement to classroom English instruction under the Kurikulum Merdeka. This requires close alignment of digital dictionary content — including vocabulary word lists, reading text genres, and contextual examples — with the thematic units and communicative functions specified in the Grade 7, 8, and 9 English curriculum. Teacher-facing features, including class progress dashboards, vocabulary mastery reports by student, and suggested in-class follow-up activities for words identified as challenging by multiple students, will support pedagogical integration and increase the likelihood of sustained use.

Theoretical Contributions

Beyond practical design implications, this systematic review makes several theoretical contributions to the literature on technology-enhanced EFL learning and self-regulated learning in digital environments. Most significantly, the review synthesizes evidence supporting a tripartite outcome model in which interactive digital dictionary use influences vocabulary mastery, reading ability, and self-regulated learning as simultaneously co-developing competencies rather than as independent outcomes of separate instructional interventions. The evidence reviewed suggests that these three outcomes are mutually reinforcing: vocabulary growth facilitated by multimodal digital dictionary features reduces the lexical barrier to reading comprehension, which in turn provides authentic contexts for deeper vocabulary learning; while SRL development supports strategic and systematic use of digital dictionary features, amplifying the vocabulary and reading gains achievable through the tool.

This tripartite model represents a departure from the prevailing tendency in CALL research to examine vocabulary, reading, and learning strategy outcomes separately, and provides a theoretically coherent framework for conceptualizing how well-designed digital learning tools can simultaneously develop multiple dimensions of language competence. The framework has implications beyond digital dictionary research, suggesting that technology-enhanced language learning tools designed with explicit attention to both the content outcomes and the metacognitive processes of language learning may produce more holistic and durable language development than tools focused exclusively on content delivery.

The review also contributes to the theoretical understanding of SRL in technology-enhanced language learning by providing evidence for the dual role of SRL as both an outcome of and a mechanism for digital tool effectiveness. This bidirectional relationship — in which digital tool use supports SRL development, and SRL in turn amplifies the benefits of digital tool use — suggests a positive developmental spiral in which learners who begin using well-designed digital dictionaries with limited SRL capacity can gradually develop stronger regulatory competencies that enable them to benefit more fully from the tool's affordances over

time. This theoretical insight has important longitudinal implications for research design and product evaluation.

CONCLUSION

This systematic literature review synthesized evidence from 45 peer-reviewed empirical studies published between 2014 and 2024 on the development and effectiveness of interactive digital dictionaries in EFL learning contexts. The review addressed five research questions examining publication trends, vocabulary mastery outcomes, reading ability outcomes, the role of self-regulated learning, and existing research gaps.

The evidence base consistently supports the effectiveness of interactive digital dictionaries as tools for enhancing vocabulary mastery across all three dimensions — breadth, depth, and fluency — with medium to large effect sizes reported across diverse EFL learner populations. Multimodal presentation of word information, contextual embedding within reading tasks, spaced repetition scheduling, and immediate corrective feedback emerge as the design features most strongly and consistently associated with vocabulary learning gains and long-term retention. Reading ability improves substantially when digital glossaries are integrated within authentic reading tasks, primarily through the mechanism of vocabulary knowledge development that raises learners above the comprehension threshold required for adequate text understanding.

Self-regulated learning emerges from this review as both a significant outcome of interactive digital dictionary use and a critical mediating mechanism through which digital dictionaries enhance vocabulary and reading outcomes. Features that explicitly support learners' goal-setting, self-monitoring, and self-evaluation — including personalized word lists, learning analytics dashboards, and adaptive review schedules — are associated with the strongest and most durable learning effects, underscoring the importance of designing digital dictionaries as tools for SRL development rather than merely as information resources.

Significant research gaps are identified, most notably the underrepresentation of Junior High School EFL learners, the scarcity of longitudinal and delayed post-test designs, the limited number of formal SRL mediation analyses, and the near-absence of studies from Southeast Asian EFL contexts including Indonesia. These gaps provide a clear and compelling rationale for the development and empirical evaluation of an interactive digital dictionary specifically designed to enhance English vocabulary mastery, reading ability, and self-regulated learning among Junior High School students in the Indonesian EFL context.

This review contributes to the field of educational technology and applied linguistics by providing a comprehensive and rigorous synthesis of existing evidence on interactive digital dictionary effectiveness, identifying the design principles most strongly supported by the empirical literature, and generating a clear research agenda for future studies. The findings provide theoretical grounding and practical guidance for the development of interactive digital dictionaries targeting the specific needs and learning contexts of Junior High School EFL learners.

REFERENCES

- Al-Jarf, R. (2021). Mobile dictionary apps and EFL vocabulary development: Self-directed learning behaviors among Saudi university students. *International Journal of English Linguistics*, 11(3), 45–58. <https://doi.org/10.5539/ijel.v11n3p45>
- Başal, A., Yilmaz, S., Tanriverdi, A., & Sari, L. (2020). Effectiveness of digital vocabulary flashcards on EFL vocabulary acquisition and retention. *Journal of Language and Linguistic Studies*, 12(2), 195–205.
- Chen, M., & Lin, H. (2022). Embedded digital annotations and reading comprehension among EFL secondary learners: A quasi-experimental study. *System*, 109, 102888. <https://doi.org/10.1016/j.system.2022.102888>
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference* (pp. 9–15). ACM.
- Ebbinghaus, H. (1885/1962). *Memory: A contribution to experimental psychology*. Dover. (Original work published 1885).
- Godwin-Jones, R. (2021). Dueling with digital dictionaries: Evolving technology for language learners. *Language Learning & Technology*, 25(1), 1–15.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>
- Khezlou, S. (2019). Form-focused instruction in CALL: What do learners think? *REL C Journal*, 50(2), 235–251. <https://doi.org/10.1177/0033688217738816>
- Lamb, R., Annetta, L., Vallett, D., Meldrum, J., & Vallett, D. (2022). Gamification in EFL vocabulary learning: Effects on motivation and vocabulary retention. *Computers & Education*, 176, 104350. <https://doi.org/10.1016/j.compedu.2021.104350>
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Macmillan.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. <https://doi.org/10.1093/applin/22.1.1>
- Lin, Y. T., Yang, S. Y., & Liang, J. C. (2022). Effects of contextual e-glossaries on EFL reading comprehension and vocabulary acquisition. *Interactive Learning Environments*, 30(4), 631–648.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nushi, M., & Eqbali, M. H. (2020). Duolingo: A mobile application to improve EFL learners' vocabulary. *Studies in Self-Access Learning Journal*, 8(3), 310–321.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson,

- E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Pham, T. N., Lin, M. S., Trinh, V. Q., & Bui, L. T. (2021). Electronic self-regulated learning and digital vocabulary learning. *Computers in Human Behavior*, 116, 106648. <https://doi.org/10.1016/j.chb.2020.106648>
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10(3), 355–371.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zhang, D., & Zou, D. (2023). Effects of digital glosses on L2 reading comprehension and vocabulary learning. *System*, 113, 102978. <https://doi.org/10.1016/j.system.2023.102978>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.