EUROPEAN JOURNAL OF EDUCATION & LANGUAGE REVIEW

European Journal of Education & Language Review, 2026, 2(1), Article 4
eISSN: 3177-7727

lectito

Conceptual paper

Multimodal Vocabulary Instruction for English Language Learners: From Interactive Whiteboards to Artificial Intelligence

Charles Martin*

Alexandria's Design LLC, UNITED STATES

***Corresponding Author:** cmartin@alexandriasdesign.com

Citation: Martin, C. (2026). Multimodal vocabulary instruction for English language learners: From interactive whiteboards to artificial intelligence. *European Journal of Education & Language Review*, *2*(1), Article 4. https://doi.org/10.20897/ejelr/18945

Published: July 9, 2026

ABSTRACT

This conceptual article examines how multimodal vocabulary instruction for English language learners (ELLs) can be designed and evaluated across successive generations of educational technology, from interactive whiteboards (IWBs) to artificial intelligence (AI)-supported tools. The article is organized around three guiding questions: what design features make technology-mediated vocabulary instruction effective for ELLs; how those features have persisted across platforms; and what affordances and risks AI-supported tools introduce. The argument is grounded in a conceptual framework that integrates research on robust vocabulary instruction, the cognitive theory of multimedia learning, and translanguaging. A qualitative doctoral study of IWB-based vocabulary instruction in a multilingual Gulf-region elementary school (Martin, 2015) is presented in detail and used as an empirical touchstone. Building on the framework and the case, the article formalizes a Design Continuity Framework comprising five platform-independent design constants (coordinated multimodal input, interactive retrieval and production, contextual embedding, actionable feedback, and spaced review) operating under two implementation conditions (teacher mediation and equitable access). Meta-analytic evidence from mobile-assisted, game-based, and chatbot-assisted vocabulary learning is then read through this framework as a cross-platform comparison. The article closes with an expanded treatment of translanguaging, a substantive discussion of vocabulary assessment, and implications for teachers (including AI literacy), instructional designers, researchers, and policymakers.

Keywords: multimodal learning, vocabulary instruction, English language learners, interactive whiteboards, educational technology, artificial intelligence, translanguaging, computer-assisted language learning

Educational technology changes quickly, but the instructional design principles that make technology useful often outlast the tools through which they were first implemented. Vocabulary instruction for English language learners (ELLs) is a clear example. Learners need repeated, meaningful, and supported encounters with new words: they need to see words, hear them, use them, connect them to context, and return to them over time (Beck et al., 2013; Nation, 2013; Schmitt, 2008). A technology is valuable to the extent that it helps teachers organize those encounters in ways that are instructionally coherent, cognitively manageable, and accessible to multilingual learners. Yet the scholarship on technology-mediated vocabulary learning is often read through the device of the moment, so that studies conducted with older platforms are filed away as relics even when their instructional logic remains current. As platforms have moved from multimedia software to interactive whiteboards (IWBs), mobile applications, digital games, virtual reality (VR), and now artificial intelligence (AI)-supported tools, the continuity of the pedagogy underneath them has become harder to see.

This article addresses that problem conceptually. It asks three guiding questions. First, *what design features make technology-mediated vocabulary instruction effective for English language learners*? Second, *how have those features persisted across technological generations, from IWBs to AI-supported tools*? Third, *what affordances and risks do AI-supported tools introduce for multimodal vocabulary instruction, and under what conditions do they strengthen rather than weaken it*? These questions are conceptual rather than empirical: the article does not report a new intervention. Instead, it develops and defends a framework through which both older and newer empirical work can be read, compared, and applied.

The argument proceeds in seven steps, and it is useful to make that structure explicit at the outset. The next section establishes the conceptual and theoretical framework, integrating research on robust vocabulary instruction, the cognitive theory of multimedia learning, and translanguaging, and defines the key terms used throughout. The third section presents the empirical touchstone for the argument: a qualitative doctoral study of IWB-based vocabulary instruction in a multilingual Gulf-region elementary classroom (Martin, 2015), including its research question, methodology, findings, and limitations. The fourth section formalizes the central conceptual contribution, a Design Continuity Framework that names the platform-independent design constants of effective technology-mediated vocabulary instruction. The fifth section reads meta-analytic and review evidence from mobile-assisted, game-based, VR-based, and chatbot-assisted vocabulary learning through that framework, providing an evidence-based cross-platform comparison. The sixth and seventh sections apply the framework to AI-supported tools and to translanguaging, and an eighth section develops the question of assessment. The article closes with implications for teachers, instructional designers, researchers, and policymakers.

The argument is intentionally conservative. It does not claim that technology automatically improves vocabulary learning. Golonka et al. (2014) caution that the evidence for technologies in foreign language learning varies considerably by technology type and research design, and that warning remains apt because current enthusiasm around AI can repeat the overclaims once made for earlier tools. The appropriate evaluative question is not whether a tool is new, but whether it realizes the design features that research identifies as instructionally valuable.

CONCEPTUAL AND THEORETICAL FRAMEWORK

Three bodies of theory and research anchor the analysis: vocabulary acquisition research, the cognitive theory of multimedia learning, and translanguaging. Together they explain why multimodal design supports vocabulary learning for ELLs, and they supply the criteria from which the Design Continuity Framework in a later section is derived. Stating these assumptions explicitly matters because the claims made throughout the article, about what vocabulary instruction requires and about how technology should be evaluated, follow from them.

Vocabulary acquisition for english language learners

Research on vocabulary instruction converges on the conclusion that word learning is incremental, multidimensional, and dependent on repeated meaningful encounters. Beck et al. (2013) argue that robust vocabulary instruction requires rich engagement with high-utility words: student-friendly explanations, multiple contexts, active processing, and cumulative review, rather than single exposures to definitions. Graves (2006) similarly describes a comprehensive vocabulary program as combining rich language experiences, direct teaching of individual words, word-learning strategies, and word consciousness. Schmitt (2008) synthesizes instructed second language vocabulary research to show that learners develop knowledge of form, meaning, use, associations, grammar, collocation, and register gradually across encounters, and Nation (2013) frames vocabulary learning as a course-design problem involving goals, deliberate learning, learning from context, strategy development, and assessment. Meta-analytic evidence supports this design-centered view: S. Webb et al. (2020) found that intentional vocabulary-learning activities produce large gains overall, but that effectiveness varies with the nature of the activity, including opportunities for retrieval and the depth of processing required, rather than with any particular delivery medium.

For English language learners, the stakes are particularly high. August et al. (2005) describe vocabulary development as critical for ELLs' reading comprehension and academic achievement, and Carlo et al. (2004) demonstrate that vocabulary instruction designed for mixed bilingual and mainstream classrooms can close part of the gap between English learners and English-only peers. Academic vocabulary can function as a gatekeeper to content learning when instruction assumes lexical resources that learners do not yet have. Multimodal design offers one way to lower that barrier by distributing access to meaning across images, speech, print, gesture, context, and interaction, an approach also reflected in sheltered instruction models such as the Sheltered Instruction Observation Protocol (SIOP), which make comprehensible input and multimodal scaffolding central to content teaching for ELLs (Echevarria et al., 2017).

The cognitive theory of multimedia learning

The cognitive theory of multimedia learning explains why coordinated modes support word learning. Mayer (2020) argues that people learn more deeply from words and pictures designed to work together than from words alone, because learners process verbal and visual information through partially separate channels with limited capacity. Mayer and Moreno (2003) warn, however, that multimedia instruction must manage cognitive load: modes that are merely added, rather than aligned with the learning task, can overload working memory and depress learning. Cognitive load theory generalizes this point, showing that instructional designs succeed when they direct limited working-memory resources toward the processing that builds knowledge and away from extraneous demands (Sweller et al., 2019). The implication for vocabulary instruction is precise: multimodality is not beneficial because it contains many modes, but because well-chosen modes clarify meaning, reduce unnecessary effort, and create multiple retrieval routes. Earlier second language multimedia research established this empirically. Chun and Plass (1996) found that multimedia annotations combining text and images supported vocabulary acquisition during reading, and Plass et al. (1998) showed that visual and verbal support in a multimedia learning environment interacted with learner preferences. These studies predate both IWBs and AI tools, yet they articulate the principle on which both depend technology-mediated vocabulary learning is strongest when word form, meaning, image, explanation, and use are connected in ways that support comprehension and retention.

Translanguaging

The third strand reframes the learner. García and Wei (2014) challenge rigid separations between named languages and describe bilingual meaning-making as the flexible deployment of a unitary linguistic repertoire. Cenoz and Gorter (2021) translate this stance into pedagogical translanguaging: planned instructional use of learners' full repertoires to advance learning in the target language. For vocabulary instruction, the implication is that learners' home languages are resources for building meaning rather than obstacles to English acquisition. A vocabulary lesson can pair an English target word with visual support, pronunciation, sentence context, and a brief cross-linguistic bridge such as a cognate comparison or home-language gloss. Within the present framework, translanguaging is treated not as an add-on but as part of multimodal design itself: the home language is one of the modes through which meaning becomes accessible. This point is developed at length in a later section, where it is connected to empirical observations from the case study.

Key terms

Because reviewing technology research across decades invites terminological drift, four terms are used in delimited senses throughout. An *interactive whiteboard* (IWB) is a large touch-sensitive display connected to a computer that allows teachers and students to control and annotate projected content by touch or stylus (Higgins et al., 2007). *Mobile-assisted vocabulary tools* are applications running on phones or tablets that organize vocabulary practice and review, most commonly flashcard, spaced-repetition, and course-style applications of the kind synthesized in mobile-assisted language learning meta-analyses (Lin & Lin, 2019; Mahdi, 2018). *AI-supported tools*, as discussed here, are primarily chatbots and tutoring systems built on large language models, such as ChatGPT and comparable systems, that generate explanations, examples, dialogue, and feedback in response to learner input (Kohnke et al., 2023; Zhang & Huang, 2024); the term AI is not used to denote a single tool but this family of generative, language-producing systems. Finally, *multimodal vocabulary instruction* refers to instruction that deliberately coordinates two or more modes (visual, auditory, textual, tactile, gestural, or interactive) around target words, in the sense developed by the multimedia learning literature reviewed above.

AN EMPIRICAL TOUCHSTONE: THE 2015 INTERACTIVE WHITEBOARD STUDY

The conceptual argument of this article grew out of, and is illustrated by, a doctoral study of IWB-based vocabulary instruction conducted in 2015 (Martin, 2015). The study is given substantive weight here because its design, findings, and limitations ground several of the claims developed later.

Context, purpose, and research question

The study took place in a private boys' elementary school in a Gulf-region city where the language of schooling was English, but the home language of nearly all students was Arabic. The local education authority had invested heavily in classroom technology, and IWBs were installed in classrooms, but teachers had little guidance on how to use them for language development. Standardized measures and authority reports at the time identified English vocabulary as a persistent area of weakness. The purpose of the study was to investigate how IWB-based materials could assist vocabulary development for third-grade ELLs, guided by a single exploratory research question: how can IWBs be used to facilitate vocabulary development in an ELL classroom?

METHODOLOGY

The study used an exploratory qualitative design rather than an experimental one; it asked how IWB-supported vocabulary instruction worked in practice, not whether it outperformed a comparison condition. Participants were two native-English-speaking teachers, each teaching two classes of approximately 25 third-grade ELLs, for a total of four classes. The researcher, embedded in the school as the grade-level planner responsible for literacy lessons, designed two IWB-based vocabulary lessons grounded in second language acquisition principles of comprehensible input and in published guidance on IWB use. The first lesson emphasized visual scaffolding, pairing high-frequency target words with images, pronunciation, and sentence contexts; the second emphasized the tactile affordances of the board, requiring students to manipulate words and images, sort items, and construct simple and complex sentences at the IWB. Both participating teachers reviewed and adapted the lessons before implementation, contributing input on oral language practice and differentiation.

Data were collected from four sources. Each teacher was interviewed three times using a semi-structured protocol, before, between, and after the observed lessons. Four classroom observations (two per teacher) were conducted using an observation tool adapted from the SIOP model and the International Society for Technology in Education classroom observation tool, attending to scaffolding, interaction with language and content, application, and practice (Echevarria et al., 2017). Student and teacher interaction with the board was captured with screen-recording software, and the researcher maintained a reflective journal to document planning conversations, observations, and potential researcher bias arising from the dual teacher-researcher role. Data were analyzed using thematic analysis: transcripts, observation notes, and journal entries were coded, codes were grouped into categories, and categories were refined into themes across the full data set.

FINDINGS

Three themes described how the IWB supported vocabulary development, and each corresponds to a modality. The tactile theme captured the value teachers and students found in physical manipulation: students came to the board to drag, sort, match, and write, and teachers reported that this manipulation increased participation and gave them immediate windows into student understanding. The auditory theme centered on repetition: hearing target words repeatedly, in the teacher's modeling, in recorded pronunciations, and in choral and individual student production, was identified by both teachers as essential for their learners. The visual theme captured the association of words with images: pairing target words with pictures and realia helped students access meaning that definitions alone did not convey.

Beyond the three modality themes, the analysis identified a set of instructional techniques that the IWB made easier to implement: pre-teaching target vocabulary; differentiating tasks within whole-class instruction; supporting bi-literacy by displaying and voicing words in both English and Arabic; teacher modeling of oral language use; structured opportunities for students to order words into simple and complex sentences; repetition across the lesson; and grouping and formative assessment strategies, with teachers reporting that the visible, manipulable record of student responses on the board helped them assess understanding and plan subsequent lessons. Two of these techniques deserve emphasis because they anticipate themes developed later in this article. The bi-literacy finding, in which students encountered target words in both English and Arabic, is an early classroom instance of what the translanguaging literature now describes systematically. The formative assessment finding anticipates the argument that multimodal vocabulary instruction must be evaluated by more than immediate recall.

Limitations and warrant

The study's limitations bound what it can support. It was small in scale, qualitative, and conducted at a single site; it examined teacher perceptions and observed practice rather than measured vocabulary gains; and the researcher's embedded role, while methodologically documented, shaped what was noticed. The study therefore cannot establish that IWBs caused vocabulary growth, and no such claim is made here. What the study can warrant, and what it contributes to this article, is a grounded account of how multimodal design principles became operationally visible in one multilingual classroom: which affordances teachers actually used, which techniques the technology made more feasible, and how modalities were coordinated in practice. Its role in the argument is illustrative and generative, not evidentiary in the experimental sense, and the framework developed next stands on the published research base even where the case adds texture.

THE DESIGN CONTINUITY FRAMEWORK

The history of educational technology in language teaching is often narrated as a sequence of discontinuous innovations: multimedia software, IWBs, mobile applications, digital games, VR, and now AI-powered tutors. Each transition tends to be framed as replacement, in which the new tool renders the old obsolete and studies conducted with older tools lose their relevance. This narrative obscures what persists. Across these generational shifts, the features that make vocabulary instruction effective do not change, because they derive from the cognitive and linguistic requirements of vocabulary acquisition rather than from any device. This section names that persistence as a framework so that it can be applied, tested, and cited: the Design Continuity Framework.

The framework holds that effective technology-mediated vocabulary instruction for English language learners realizes five platform-independent design constants, operating under two implementation conditions. The five constants are derived directly from the theoretical framework established earlier. First, coordinated multimodal input: word form, pronunciation, image, and context are presented in deliberate combination, consistent with multimedia learning theory (Mayer, 2020) and with the visual and auditory themes of the case study. Second, interactive retrieval and production: learners must do something with words (retrieve, discriminate, produce, apply) rather than only receive them, consistent with vocabulary research showing that retrieval and depth of processing drive learning (Schmitt, 2008; S. Webb et al., 2020) and with the tactile theme of the case. Third, contextual embedding: words are met in sentences, tasks, and content, not only in lists, consistent with robust vocabulary instruction (Beck et al., 2013). Fourth, actionable feedback: learners receive responses that help them refine form and meaning, and teachers can see and interpret learner performance. Fifth, spaced review: instruction returns to words across lessons and contexts, consistent with the incremental nature of word learning (Nation, 2013).

The two implementation conditions are not modalities but preconditions that determine whether the constants operate as intended. Teacher mediation refers to the teacher's role in selecting words, sequencing encounters, interpreting feedback, and correcting tool output; the case study found that lessons not designed or adapted by the teacher required adaptation to the specific classroom before they worked, and the AI literature reviewed below reinforces the point (Kohnke et al., 2023; Zawacki-Richter et al., 2019). Equitable access refers to whether all learners, including those in low-resource settings and those whose home languages are less well served by commercial tools, can actually reach the instruction. Table 1 (below) summarizes the framework.

Table 1

The design continuity framework: Constants, rationale, and platform expressions

Component	Theoretical basis	Example expressions across platforms
Coordinated multimodal input	Multimedia learning theory (Mayer, 2020; Mayer & Moreno, 2003)	IWB image-word-audio displays; app flashcards with audio; chatbot example sentences with images
Interactive retrieval and production	Retrieval and depth of processing (Schmitt, 2008; S. Webb et al., 2020)	Tactile sorting at the board; app recall prompts; game tasks; chatbot dialogue requiring word use
Contextual embedding	Robust vocabulary instruction (Beck et al., 2013)	Sentence construction at the IWB; game narratives; VR scenes; chatbot-generated contexts
Actionable feedback	Formative assessment in vocabulary learning (Nation, 2013)	Visible student responses on the board; app correctness feedback; chatbot corrective feedback
Spaced review	Incremental vocabulary growth (Nation, 2013; Schmitt, 2008)	Cumulative IWB review slides; spaced-repetition scheduling; chatbot review conversations
Condition A: Teacher mediation	Educator-centered technology research (Zawacki-Richter et al., 2019)	Teacher adaptation of lessons; review of AI output; interpretation of learner data
Condition B: Equitable access	Equity in language education (August et al., 2005; García & Wei, 2014)	Low-tech analogues; device and connectivity availability; home-language quality of AI output

Three clarifications protect the framework from misreading. First, the constants are necessary considerations, not a guarantee: a design can include all five and still fail through poor word selection or excessive cognitive load. Second, the framework prevents technological determinism in both directions. An IWB does not guarantee multimodal learning, and AI does not guarantee personalization; a teacher can use an IWB for passive lecture, and a student can use a chatbot for shallow answer generation. Conversely, a printed image set, a teacher's voice, peer discussion, and a gesture routine can satisfy every constant without any digital tool, which is why the framework supports rather than undermines low-resource classrooms. Third, the framework changes how research should be read: a study's durable contribution lies in what it shows about the constants, not in the platform it happened to use. The next section applies that reading to the empirical literature.

READING THE EVIDENCE ACROSS PLATFORMS

If the Design Continuity Framework is sound, then the empirical literature on technology-assisted vocabulary learning should show two patterns: positive but variable effects within every platform generation, and moderation of those effects by design features rather than by platform novelty. Both patterns appear, and reading the meta-analytic record through the framework provides the evidence-based cross-platform comparison that a purely narrative history cannot.

At the broadest level, Yu and Trainin (2022) meta-analyzed technology-assisted second language vocabulary learning across tool types and found a moderate overall advantage for technology-assisted conditions, with significant variability across studies; critically for the present argument, their moderator analyses indicated that outcomes differed by instructional arrangement and learning condition rather than tracking any simple hierarchy of tool recency. Simonnet et al. (2025), reviewing technology-assisted vocabulary learning from 2015 to 2023, reached a complementary conclusion by shifting the unit of analysis from technology categories to the instructional features embedded in vocabulary tools, an analytic move that the Design Continuity Framework formalizes. S. Webb et al. (2020), although focused on intentional vocabulary-learning activities rather than technology as such, supply the activity-level baseline: gains follow from what learners are required to do with words, which is precisely what the framework's second constant predicts.

Within the mobile generation, Mahdi (2018) found a positive overall effect of mobile devices on vocabulary learning, with moderator analyses showing stronger effects under some conditions of use than others. Lin and Lin (2019), synthesizing mobile-assisted ESL and EFL vocabulary studies, likewise reported positive effects moderated by intervention design features such as duration and the kind of practice involved. Duan et al. (2026), reviewing mobile-assisted English vocabulary learning from 2019 to 2025, found consistent benefits for retention and motivation while identifying the same gaps the framework highlights: insufficient attention to deeper word knowledge, contextual use, and teacher facilitation. Read together, the mobile literature shows constants one, two, and five operating (multimodal flashcards, retrieval practice, spaced scheduling) while flagging weaknesses in constants three and four and in the teacher-mediation condition.

The game-based and immersive literatures extend the pattern. Zou et al. (2021) reviewed digital game-based vocabulary learning and found generally positive outcomes, with effectiveness linked to whether games kept lexical processing central to play, contextual embedding and retrieval, rather than to engagement alone. Tai et al. (2022) found that a VR application improved adolescent EFL learners' vocabulary learning relative to a conventional condition, attributing the advantage to contextualized, interactive encounters with words in simulated environments, an intensification of constant three. In the chatbot generation, Wang et al. (2025) meta-analyzed chatbot-assisted language learning and found positive overall effects with moderators including educational level, interface design, and interaction capability, and Zhang and Huang (2024) reported vocabulary-specific gains from large language model chatbots. In every generation, in other words, the platforms that realized more of the constants produced better vocabulary outcomes, and the moderator analyses repeatedly locate effectiveness in design features the framework names rather than in the platform category itself. This convergence across independently conducted syntheses, spanning two decades and at least four platform generations, is the strongest empirical warrant for design continuity as an organizing principle.

AI-SUPPORTED TOOLS: AFFORDANCES AND LIMITATIONS

AI-supported tools are the current platform through which the framework's questions are being asked again—research attention to AI in education has expanded rapidly over the past decade, with personalized learning and, most recently, generative AI as dominant themes (Almarashdi et al., 2026)—and they deserve focused treatment both because their affordances are real and because their risks are distinctive. On the affordance side, large language model chatbots can generate unlimited example sentences at controlled difficulty, adjust explanations to a learner's level, simulate conversation that requires target-word production, and provide immediate feedback, capabilities that map directly onto constants one through four and that earlier platforms supplied only partially (Kohnke et al., 2023; Zhang & Huang, 2024). When combined with speech, image, or translation features, a single AI-supported interaction can approximate the coordinated multimodal episode that the IWB study realized through whole-class instruction, but at an individualized scale: a learner can see an image, hear a word, request simpler examples, compare near-synonyms, produce a sentence, and receive correction within one exchange.

The limitations are equally structural. Kasneci et al. (2023) catalogue the risks of large language models in education, including fluent but inaccurate output, bias, and opacity, and Zawacki-Richter et al. (2019) document a longer-standing pattern in which AI-in-education research neglects educators, treating teachers as bypassable rather than as the agents who make tools instructionally meaningful. For vocabulary instruction specifically, a chatbot can generate definitions that are wrong, examples that are too advanced or culturally inappropriate, and feedback that appears authoritative even when it is mistaken; young learners are poorly positioned to evaluate such output without teacher mediation. Privacy and access compound the problem: tools that collect children's voice, writing, and performance data require governance, and if high-quality AI support depends on subscriptions, newer devices, or stable broadband, it can widen the very gaps that multilingual instruction is meant to reduce—digital inequality operates through learners' whole socio-technical circumstances, not device access alone (Katz, 2026). Output quality also varies across languages and dialects, which matters directly for the translanguaging affordances discussed next.

These considerations motivate an explicit instructional response rather than avoidance: AI literacy, for teachers and, in age-appropriate forms, for learners. AI literacy is conceptualized as the knowledge and skills needed to understand, use, evaluate, and critically appraise AI systems (Ng et al., 2021), and international guidance now recommends building such capacity in teachers as a condition of responsible classroom use of generative AI (UNESCO, 2023). Recent evidence from teacher preparation underscores the need: pre-service teachers working with multilingual learners report difficulty recognizing subtle media bias and evaluating AI-amplified misinformation, and limited confidence in translating that awareness into instructional strategies (Halpern et al., 2025). Within the Design Continuity Framework, AI literacy is the form that the teacher-mediation condition takes when the platform generates language: teachers need to know when to trust, verify, correct, and discard AI output, and learners need to understand that fluent text is not therefore accurate text. The implications section returns to this point as a concrete recommendation, because a balanced position on AI is not a vague middle ground but a set of teachable evaluative practices.

TRANSLANGUAGING AS MULTIMODAL DESIGN

Translanguaging deserves fuller integration with the multimodal framework than it usually receives, because the theoretical fit is exact: if multimodal design works by giving learners multiple coordinated routes into word meaning, then the home language is a mode, arguably the most powerful one a multilingual learner brings. García and Wei (2014) establish the descriptive claim that bilingual learners deploy a unitary repertoire, and Cenoz and Gorter (2021) establish the pedagogical claim that instruction can plan for that repertoire deliberately rather than merely tolerating it. Within the Design Continuity Framework, translanguaging strengthens constant one, because a home-language gloss or cognate comparison is coordinated input that clarifies meaning at low cognitive cost; constant three, because cross-linguistic connections embed new words in conceptual knowledge the learner already holds; and condition B, because instruction that builds on home languages is accessible to learners whom English-only design excludes.

The case study offers a concrete classroom image of this integration. In the observed lessons, target words were displayed and voiced in both English and Arabic, and teachers reported that this bi-literacy support helped students access meaning and increased participation (Martin, 2015). The practice was unremarkable in the classroom and entirely consistent with the school's English-medium goals: the home language served as a bridge into English production, not a substitute for it. AI-supported tools now make such bridges trivially easy to generate (a chatbot can produce a translation, cognate comparison, or simple home-language sentence on demand), but the quality caveat raised above applies with force, because output quality varies across languages, and inaccurate equivalents can flatten nuanced meanings. The design implication is that translanguaging supports should be planned, reviewed, and connected to English production, with the teacher, or for older learners the students themselves, positioned to evaluate machine-generated cross-linguistic content. Treated this way, translanguaging is not a separate equity gesture appended to multimodal instruction; it is one of the modes the framework coordinates, and equity is foregrounded as a property of the design itself.

ASSESSING MULTIMODAL VOCABULARY LEARNING

A framework that specifies what vocabulary instruction should do must also specify how learning should be judged, and assessment is where technology-assisted vocabulary research has been weakest. Most studies in the syntheses reviewed above measure recognition or short-term recall immediately after instruction, yet vocabulary knowledge is multidimensional, spanning form, meaning, and use, receptive and productive mastery, and breadth and depth (Nation, 2013; Read, 2000; Schmitt, 2008). A learner who matches a word to a picture immediately after a lesson may still be unable to use the word in an explanation, recognize it in a new text, or connect it to related words; conversely, instruments that capture only immediate recognition will systematically overstate the effectiveness of designs that drill recognition and understate designs that build productive, transferable knowledge. S. Webb et al. (2020) document exactly this sensitivity: measured effectiveness varies with the match between the learning activity and the test format.

Three assessment commitments follow for multimodal vocabulary instruction. First, assessment should sample more than one dimension of word knowledge, pairing recognition measures with productive tasks such as sentence construction or oral explanation, the very tasks the case-study teachers used informally at the IWB. Second, assessment should include delayed measures, because spaced review (constant five) is justified precisely by its effect on retention, which immediate post-tests cannot show. Third, formative assessment should be treated as a design feature rather than an afterthought: the case study found that the visible, manipulable record of student responses on the board functioned as continuous formative evidence that teachers used to plan subsequent lessons (Martin, 2015), and AI-supported tools could extend this affordance by logging and summarizing learner production, provided the feedback remains interpretable by teachers (constant four and condition A). Researchers evaluating new platforms should report which dimensions of word knowledge their measures capture and at what delay; practitioners selecting tools should ask whether a tool's built-in measures assess anything beyond recognition. An evaluative framework of this kind keeps assessment aligned with the construct (vocabulary knowledge that transfers into listening, speaking, reading, writing, and content learning) rather than with what is easiest to measure.

IMPLICATIONS

For Teachers

Six design practices follow from the framework and the evidence. First, select words that matter for comprehension and academic participation, consistent with robust vocabulary instruction (Beck et al., 2013; Graves, 2006). Second, introduce words through coordinated modes that clarify meaning without overloading learners (Mayer, 2020; Mayer & Moreno, 2003; Sweller et al., 2019). Third, require active retrieval and production, because word learning follows from what learners do with words (Schmitt, 2008; S. Webb et al., 2020). Fourth, build review across days and contexts, using mobile or digital tools when they make distributed practice easier (Lin & Lin, 2019). Fifth, treat learners' home languages as planned resources, using brief cross-linguistic bridges connected to English production (Cenoz & Gorter, 2021; García & Wei, 2014). Sixth, develop AI literacy: when AI-supported tools enter the classroom, teachers should verify generated definitions and examples before or during instructional use, teach learners age-appropriate skepticism toward fluent machine output, and prefer tools whose feedback they can inspect and interpret (Ng et al., 2021; UNESCO, 2023). None of these practices requires advanced technical expertise; all of them require clarity about why multimodal input, interaction, feedback, and review support vocabulary learning. Global evidence on English language teacher development supports this emphasis on pedagogy over platform: across 43 countries, teachers' core instructional beliefs remained stable through the pandemic's forced technology shift, while digital infrastructure shaped only how those beliefs were enacted (R. Webb, 2026). A teacher without tablets, an IWB, or AI tools can satisfy every constant with printed images, oral pronunciation, student-friendly explanations, gesture, peer discussion, and planned review, which is why the framework serves low-resource classrooms rather than excluding them.

For instructional designers

Designers working at scale should treat the five constants as design requirements and the two conditions as review criteria. A vocabulary screen should not simply list a word, a definition, and a stock image; it should show the word in context, provide pronunciation or listening support, require the learner to do something with the word, and return to the word later. Design reviews can operationalize the framework as an audit: For each target word, what makes it instructionally important? How will learners first access its meaning? How will they hear and pronounce it? Where will they meet it in context? What interaction requires retrieval or production? What feedback will learners and teachers see? When will the word return? How will multilingual learners connect it to existing linguistic knowledge? If AI generation is included, it should serve a specified instructional purpose (additional examples, differentiated practice, inspectable feedback) rather than functioning as a decorative layer, and generated content that reaches young learners should pass through human review.

For researchers

The framework suggests a methodological agenda. First, design-centered synthesis: vocabulary studies should be coded for the presence of the five constants and two conditions (visual and auditory support, contextualized examples, learner production, feedback, review schedule, teacher mediation, multilingual resources) so that studies from different platform generations can be compared on common dimensions, a direction in which recent feature-oriented reviews have already moved (Simonnet et al., 2025). Such coding would help identify which features are most consistently associated with retention and productive use, and would prevent the field from rediscovering the same principles under successive technological names. Second, platform comparisons that hold design constant: studies comparing IWB, mobile, game, VR, chatbot, and low-tech multimodal instruction using the same target words, populations, review schedules, and assessments would isolate whether outcomes are driven by platforms or by the design features platforms embed. Third, assessment upgrading: future studies should report delayed and productive measures alongside immediate recognition, for the reasons developed above. Finally, older studies, including small qualitative ones such as the case revisited here, should be read for what they reveal about the constants in practice rather than dismissed with their hardware; the bi-literacy and formative-assessment observations of the 2015 study remain legible and useful precisely because they are read at the level of design.

For policymakers and curriculum leaders

Technology investment decisions should be guided by pedagogical criteria rather than device novelty. Procurement of AI tutoring systems, mobile applications, VR platforms, or interactive displays should include an instructional design audit against the constants and conditions: coordinated multimodal input, interactive learner response, contextual embedding, interpretable feedback, structured review, teacher mediation, data privacy, and equitable access, including the quality of support for learners' home languages. Evidence requirements should extend beyond engagement and usage metrics to vocabulary learning, retention, productive use, and teacher usability across learner groups. Finally, multilingual supports should be built into core vocabulary instruction from the beginning rather than appended as remedial add-ons; visuals, oral models, home-language bridges, sentence frames, interactive retrieval, and spaced review benefit all learners while reducing stigma for those acquiring English.

CONCLUSION

This article has argued that the movement from interactive whiteboards to artificial intelligence in vocabulary instruction is best understood not as a story of replacement but as a demonstration of design continuity. The learning problem is stable: English language learners need repeated, meaningful, multimodal encounters with academic vocabulary, encounters that coordinate input, require retrieval and production, embed words in context, return feedback, and revisit words over time. The Design Continuity Framework names these constants and the conditions (teacher mediation and equitable access) under which they operate, and the meta-analytic record across mobile, game-based, immersive, and chatbot platforms supports reading effectiveness at the level of design features rather than device categories. The 2015 IWB study revisited here illustrates how those features looked in one multilingual classroom, including early classroom expressions of translanguaging and formative assessment that current tools can now extend at scale. AI-supported tools will reward the same design discipline as their predecessors and punish its absence, while adding risks of inaccuracy, opacity, and inequitable access that make teacher mediation, and the AI literacy it now requires, part of the instructional design itself. The most useful evaluative questions for the next platform generation are therefore already available: Does the design engage multiple modes purposefully? Does it require the learner to act, not only receive? Does it return to words over time? Does it provide feedback that teachers and learners can understand? Does it open rather than close access for multilingual learners? Tools that answer those questions affirmatively will extend the promise of multimodal vocabulary instruction rather than merely rebranding it.

Acknowledgement

The author thanks the editor and the anonymous reviewers for their constructive feedback on earlier versions of this article. The author is also grateful to his wife, Marie, for her steady support and patience throughout this work, and to his parents and brother for a lifetime of encouragement that made this path possible. Finally, the author dedicates this work to his daughter, whose curiosity is a constant reminder of why learning matters.

Funding

This work received no external funding.

Ethical statement

The empirical study informing this article was conducted as part of the author’s doctoral research at the University of Florida and received approval from the University of Florida Institutional Review Board prior to data collection. Permission to conduct the study was obtained from the participating school’s administration, and informed consent was obtained from participating adults and from the parents or guardians of participating students, with student assent, in accordance with the approved protocol. All identifying information about the school, teachers, and students has been withheld to protect participant confidentiality. No new human-subjects data were collected for the present article.

Competing interests

The author declares no competing interests.

Author contributions

Charles Martin is the sole author of this article, and is responsible for all aspects of it, including conceptualization; Writing (Original draft, review & editing) and Formal Analysis (synthesis of theoretical and empirical literature).

Data availability

The empirical touchstone described in this article draws on the author's doctoral dissertation (Martin, 2015), which is publicly available. No new datasets were generated or analyzed for the present article.

AI disclosure

AI-assisted tools were used for editorial organization, language polishing, formatting checks, and citation-integrity review. All substantive claims, analysis, source selection, and final text were reviewed and approved by the author, who takes full responsibility for the manuscript.

Biographical sketch

Charles Martin, Ed.D., is the founder of Alexandria's Design LLC, a learning design studio in the United States. He holds a Doctor of Education in educational technology from the University of Florida, where his doctoral research examined interactive whiteboard-supported vocabulary instruction for third-grade English language learners in the United Arab Emirates. A former elementary teacher and school administrator, including in international English-medium schools, he has designed digital learning experiences for K-12, higher education, government, and corporate settings, and taught as a university adjunct instructor. His work focuses on multimodal vocabulary instruction, multilingual learners, and artificial intelligence in language education.

Disclaimer/Publisher's Note

REFERENCES

Almarashdi, H. S., Abu Khurma, O., AlArabi, K., Abulebdeh, E., & Yousef, J. (2026). AI-enhanced STEM education: A bibliometric study of research trends toward achieving sustainable development goals. *European Journal of STEM Education, 11*(1), 16. https://doi.org/10.20897/ejsteme/18190

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50–57. https://doi.org/10.1111/j.1540-5826.2005.00120.x

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188–215. https://doi.org/10.1598/RRQ.39.2.3

Cenoz, J., & Gorter, D. (2021). *Pedagogical translanguaging*. Cambridge University Press. https://doi.org/10.1017/9781009029384

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal, 80*(2), 183–198. https://doi.org/10.1111/j.1540-4781.1996.tb01159.x

Duan, H., Yunus, M. M., & Ismail, H. H. (2026). Mobile-assisted English vocabulary learning: A systematic review of platforms, learning outcomes, and implications (2019–2025). *International Journal of Learning, Teaching and Educational Research, 25*(2), 65–90. https://doi.org/10.26803/ijlter.25.2.4

Echevarria, J., Vogt, M. E., & Short, D. J. (2017). *Making content comprehensible for English learners: The SIOP model* (5th ed.). Pearson.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Pivot. https://doi.org/10.1057/9781137385765

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. Teachers College Press.

Halpern, B., Aydin, H., & Halpern, C. (2025). Seeing multilingual learners through media and AI: Pre-service teachers' perceptions in an ESOL course. *Journal of Interdisciplinary Research in Artificial Intelligence and Society, 1*(1), 4. https://doi.org/10.20897/jirais/17647

Higgins, S., Beauchamp, G., & Miller, D. (2007). Reviewing the literature on interactive whiteboards. *Learning, Media and Technology, 32*(3), 213–225. https://doi.org/10.1080/17439880701511040

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Katz, N. (2026). Digital inequality in context: A socio-technical analysis of Arab students' remote learning in Israel. *Asia Pacific Journal of Education and Society, 14*(1), 6. https://doi.org/10.20897/apjes/17975

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lin, J.-J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Computer Assisted Language Learning, 32*(8), 878–919. https://doi.org/10.1080/09588221.2018.1541359

Mahdi, H. S. (2018). Effectiveness of mobile devices on vocabulary learning: A meta-analysis. *Journal of Educational Computing Research, 56*(1), 134–154. https://doi.org/10.1177/0735633117698826

Martin, C., Jr. (2015). *Teaching vocabulary development using interactive whiteboards to third grade ELL students in the United Arab Emirates* [Doctoral dissertation, University of Florida].

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781316941355

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139858656

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25–36. https://doi.org/10.1037/0022-0663.90.1.25

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732942

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Simonnet, E., Loiseau, M., & Lavoué, É. (2025). A systematic literature review of technology-assisted vocabulary learning. *Journal of Computer Assisted Learning, 41*(1), e13096. https://doi.org/10.1111/jcal.13096

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning, 35*(4), 892–917. https://doi.org/10.1080/09588221.2020.1752735

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Wang, F., Cheung, A. C. K., Neitzel, A. J., & Chai, C. S. (2025). Does chatting with chatbots improve language learning performance? A meta-analysis of chatbot-assisted language learning. *Review of Educational Research, 95*(4), 623–660. https://doi.org/10.3102/00346543241255621

Webb, R. (2026). Post pandemic English language teacher development: A global perspective. *European Journal of Education & Language Review, 2*(1), 1. https://doi.org/10.20897/ejelr/17719

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *The Modern Language Journal, 104*(4), 715–738. https://doi.org/10.1111/modl.12671

Yu, A., & Trainin, G. (2022). A meta-analysis examining technology-assisted L2 vocabulary learning. *ReCALL, 34*(2), 235–252. https://doi.org/10.1017/S0958344021000239

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon, 10*(3), e25370. https://doi.org/10.1016/j.heliyon.2024.e25370

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: Where are we and where are we going? *Computer Assisted Language Learning, 34*(5–6), 751–777. https://doi.org/10.1080/09588221.2019.1640745

