



## Video-based learning in EFL contexts: a systematic review of empirical evidence and pedagogical implications

Sylvia

<sup>a\*</sup> Master of English Education, Faculty of Teacher Training and Education, Universitas Terbuka, Tangerang Selatan, Indonesia

\*Correspondence: [sylvia89@ecampus.ut.ac.id](mailto:sylvia89@ecampus.ut.ac.id)

Receive: 27/07/2025

Accepted: 01/08/2025

Published: 01/10/2025

### Abstrak

Peningkatan integrasi pembelajaran berbasis video dalam pendidikan Bahasa Inggris sebagai Bahasa Asing (EFL) telah mengubah cara peserta didik berinteraksi dengan bahasa, cara guru memfasilitasi pembelajaran, serta cara lembaga pendidikan merancang kurikulum. Tinjauan sistematis ini mensintesis berbagai studi empiris yang diterbitkan antara tahun 2023 hingga 2025, yang meneliti potensi pedagogis dan tantangan dari berbagai pendekatan berbasis video, termasuk komunikasi video sinkron, observasi yang diperkaya video, translanguaging dalam penyuntingan sejawat, teknik *shadowing* berbasis video, aplikasi dubbing, serta integrasi media sosial seperti TikTok.

Hasil penelitian menunjukkan bahwa intervensi berbasis video secara signifikan meningkatkan kemampuan berbicara, pelafalan, praktik reflektif, keterlibatan kognitif, dan keterlibatan sosial-emosional peserta didik. Namun, efektivitasnya sangat bergantung pada desain pembelajaran, tingkat kemahiran peserta didik, akses terhadap teknologi, dan dukungan institusional. Meskipun pembelajaran tatap muka cenderung memfasilitasi dinamika interaksi dan keterlibatan emosional yang lebih kaya, lingkungan berbasis video menawarkan peluang unik untuk praktik reflektif, pemrosesan input multimodal, dan pemecahan masalah secara kolaboratif.

Tantangan seperti kelebihan beban kognitif, berkurangnya kehadiran sosial, gangguan digital, serta ketimpangan akses teknologi menekankan perlunya pemberian *scaffolding* pedagogis dan integrasi yang kritis. Secara keseluruhan, tinjauan ini menyoroti peran transformatif namun sensitif terhadap konteks dari pembelajaran berbasis video dalam pendidikan EFL, serta menyerukan penelitian lebih lanjut mengenai dampak jangka panjang, strategi pembelajaran yang terdiferensiasi, dan integrasi teknologi adaptif berbasis AI untuk mengoptimalkan ekologi pembelajaran multimodal.

**Kata kunci:** pembelajaran berbasis video; pendidikan EFL; pembelajaran multimodal; alat digital

### Abstract

The increasing integration of video-based learning in English as a Foreign Language (EFL) education has transformed the ways learners engage with language, teachers facilitate instruction, and institutions design curricula. This systematic review synthesizes empirical studies published between 2023 and 2025 that examine the pedagogical affordances and challenges of diverse video-mediated approaches, including synchronous video communication, video-enhanced observation, translanguaging in peer editing, video-based shadowing, dubbing applications, and social media integration such as TikTok. Findings reveal that video-based interventions significantly enhance learners' oral proficiency, pronunciation, reflective practices, cognitive engagement, and socio-emotional involvement. However, their effectiveness is highly contingent upon instructional design, learners' proficiency levels, technological accessibility, and institutional support. While face-to-face modalities tend to foster richer interactional dynamics and emotional engagement, video-mediated environments provide unique opportunities for reflective practice, multimodal input, and collaborative problem-solving. Challenges such as cognitive overload, reduced social presence, digital distractions, and unequal access underscore the need for pedagogical scaffolding and critical integration. Overall, this review highlights the transformative yet context-sensitive role of video-based learning in EFL education, calling for future research on long-term impacts, differentiated instructional strategies, and the integration of AI-driven adaptive technologies to optimize multimodal learning ecologies.

**Keywords:** Video-based learning; EFL education; multimodal learning; digital tools

## Introduction

The rapid advancement of digital technologies has profoundly transformed the landscape of English as a Foreign Language (EFL) education. In this evolving environment, video-based learning has gained significant attention as a dynamic and multimodal approach to language instruction (Grandon, 2018). Unlike traditional text-based or audio-centered practices, the integration of video enables learners to observe authentic linguistic interactions, contextualize meaning through visual cues, and engage with real-world communication patterns (Brewin, 2025). This shift reflects a growing recognition that language learning extends beyond grammatical accuracy toward communicative competence, intercultural understanding, and learner autonomy (Avsheniuk et al., 2023).

Video-based learning offers unique pedagogical affordances that align closely with the cognitive and affective processes involved in language acquisition (Khalima & Teacher, 2025). Through multimodal exposure, learners can connect auditory and visual stimuli, facilitating deeper comprehension and retention. Visual representation of gestures, facial expressions, and situational contexts enhances learners' pragmatic awareness and supports the internalization of language use in authentic settings. Moreover, video materials can promote self-paced and reflective learning, allowing students to pause, review, and analyze linguistic input at their own convenience (Palaiageorgiou & Papadopoulou, 2019).

From an instructional perspective, video-based approaches foster active engagement and collaborative participation. Learners are not passive recipients of information but become co-constructors of meaning as they interact with visual and auditory content (Choi & Nunan, 2022). Teachers, in turn, can design video-mediated tasks that encourage observation, imitation, and creative production, thereby bridging receptive and productive language skills. Such integration can transform classrooms into interactive, learner-centered environments where communication and critical thinking naturally emerge through visual storytelling and discussion (Oyelana et al., 2022).

Equally important, video-based learning nurtures learners' socio-emotional

development alongside linguistic competence. Exposure to authentic contexts, cultures, and communicative situations fosters empathy, intercultural awareness, and motivation (Jain et al., 2024). Learners develop a stronger sense of agency as they see language as a tool for real communication rather than a set of isolated rules. The affective benefits of video-based learning—such as enjoyment, curiosity, and confidence—play a crucial role in sustaining long-term engagement, especially in remote or blended learning environments (Nowak & Kaplan-Rakowski, 2024).

However, despite its pedagogical promise, video-based learning presents several challenges that educators must navigate. Without thoughtful design, it can lead to cognitive overload or superficial engagement, where learners focus on visual appeal rather than linguistic substance (Nelson & Erlandson, 2008). Issues such as unequal digital access, limited technical literacy, and distraction risk undermining the learning process. Furthermore, ethical considerations surrounding copyright, data privacy, and online behavior must be addressed to ensure responsible and equitable implementation of video technologies in educational settings (Li & Zhang, 2025).

Given these opportunities and challenges, a comprehensive understanding of video-based learning is essential for optimizing its role in EFL instruction. It is not merely the inclusion of videos that determines success, but how these resources are integrated within coherent pedagogical frameworks that promote interaction, reflection, and meaningful communication (Smyth, 2011). By exploring the theoretical and practical implications of video-based learning, educators and policymakers can design strategies that balance innovation with inclusivity, ensuring that technology serves as a bridge—rather than a barrier—to effective language education.

## Method

This study adopts a qualitative research design grounded in the principles of systematic literature review, guided by the methodological framework proposed by Duque et al. (2024). The review aims to critically synthesize empirical evidence concerning the

implementation and pedagogical implications of video-based learning within English as a Foreign Language (EFL) contexts. A comprehensive and structured search was carried out across reputable academic databases, including Scopus, Copernicus, and Google Scholar, to ensure a wide yet methodologically consistent coverage of relevant literature.

The selection and screening process followed predefined inclusion and exclusion criteria emphasizing conceptual alignment, methodological rigor, and pedagogical significance. To maintain contemporaneity and relevance, only studies published between 2023 and 2025 were included. The search strategy employed systematically constructed keyword strings such as “*video-based learning*,” “*video-assisted instruction*,” “*video-mediated communication*,” and “*technology-enhanced language learning*.” Boolean operators (AND/OR) were applied to maximize both precision and comprehensiveness in retrieving pertinent sources.

Data extraction and analysis were conducted through a multi-stage coding process, involving identification of key pedagogical themes, instructional models, and reported learning outcomes. Thematic synthesis was employed to organize findings across cognitive, linguistic, and affective dimensions of EFL instruction. This analytical approach enabled the identification of convergent trends, contextual variations, and critical gaps in the current body of research, providing an integrative perspective on how video-mediated pedagogies inform effective, inclusive, and reflective language teaching practices.

### **Instrument and Procedures**

The principal research instrument adopted in this study was a systematically developed document analysis protocol, meticulously designed to ensure methodological rigor and analytical transparency in the extraction, classification, and synthesis of pertinent data from the reviewed literature. The review process unfolded through a series of structured, iterative, and interdependent stages to uphold the validity and reliability of the synthesis. In the preliminary phase, scholarly works

explicitly addressing *Video-Based Learning in EFL Contexts* were identified and subjected to a rigorous screening process based on predefined inclusion and exclusion criteria. Each eligible publication underwent an intensive analytical reading to extract essential information encompassing research objectives, methodological orientations, participant characteristics, instructional designs, and reported pedagogical outcomes.

Following data extraction, the collected materials were systematically organized and synthesized into six overarching thematic domains to facilitate analytical coherence and interpretive depth. These domains encapsulate the multidimensional nature of video-based learning in EFL instruction and serve as the foundation for subsequent theoretical and pedagogical analysis:

1. Pedagogical Applications – encompassing diverse instructional approaches, classroom practices, and technological affordances underpinning the implementation of video-based learning within EFL contexts;
2. Emerging Developments – capturing recent innovations, adaptive technologies, and evolving paradigms that characterize the design, integration, and evolution of video-mediated instruction;
3. Empirical Findings – consolidating evidence on the pedagogical impact of video-based interventions on learners’ linguistic proficiency, communicative competence, engagement, and critical reflection;
4. Research Gaps – identifying deficiencies in the existing literature, including the paucity of longitudinal investigations, limited cross-contextual validation, and insufficient exploration of affective, intercultural, and sociocognitive dimensions;
5. Theoretical and Pedagogical Implications – delineating conceptual contributions to contemporary language pedagogy and underscoring the transformative potential of multimodal, video-enhanced instructional frameworks; and
6. Future Research Directions – highlighting the necessity for

empirical expansion through classroom-based experimentation, longitudinal inquiry, and theory-driven pedagogical modeling to optimize the integration of video-based learning in EFL education.

This systematic and recursive procedure ensured a comprehensive, credible, and critically informed synthesis of the empirical landscape surrounding video-based learning in EFL contexts. The methodological rigor embedded in each stage of the review process reinforces the validity of the ensuing analytical interpretations and pedagogical implications derived from the corpus of contemporary scholarship.

### Result and Discussion

Recent advances in technology-enhanced language learning have foregrounded the role of digital platforms and multimodal tools in shaping learners' engagement, reflective practices, and linguistic development in EFL contexts. A growing body of research demonstrates how communication modalities, digital observation tools, translanguaging practices, video-based activities, and social media applications collectively transform task performance, pronunciation, oral proficiency, and academic competencies. These studies highlight not only the cognitive and linguistic benefits of technology integration but also the socio-emotional, reflective, and motivational dimensions that underpin effective learning experiences. Building upon these insights, the following section synthesizes empirical findings that illuminate the pedagogical affordances and challenges of diverse technology-mediated interventions in EFL education. Building on these insights, the following section outlines the key findings and discussions.

Tabel 1. Research Results

Article Title (Author)	Research Focus
Second Language Task Engagement in Face-To-Face and Synchronous Video-Based Computer-Mediated Communication	The purpose of this study is to examine differences in EFL learners' engagement across behavioral, cognitive, social, and emotional

Modes: Performances and Perceptions (Aubrey & Philpott, 2023)	dimensions when performing decision-making tasks in synchronous video-based computer-mediated communication (SvCMC) and face-to-face (FTF) modes.
The Use of Video Enhanced Observation in Video-Mediated Post-Observation Conversations on Pre-Service EFL Teachers' Online Practicum Teaching (Kanat Mutluoğlu & Balaman, 2023)	The purpose of this study is to explore how video-mediated post-observation conversations, supported by a digital observation and video tagging tool, facilitate reflective practice and professional development for pre-service EFL teachers in online practicum teaching contexts.
The Presence of Translanguaging and Its Effect on EFL Learners' Linguaging Opportunities in Video-Based Peer Editing (Alqefari, 2024)	This study aims to examine how translanguaging occurs in video-mediated peer editing among Saudi EFL learners and its impact on the production and resolution of language-related episodes (LREs).
Effects of Video-based Shadowing on Suprasegmental Features: EFL Learners' Pronunciation Performance and Attitudes (Van et al., 2024)	The study aimed to investigate the effects of video-based shadowing on pre-intermediate EFL learners' pronunciation of suprasegmental features and to explore their attitudes toward the technique.

<p>The Impact of Video Dubbing App on Chinese College Students' Oral Language Skills Across Different Proficiency Levels (Wang &amp; Lee, 2025)</p>	<p>The study aimed to examine the effectiveness of an English video dubbing app in improving Chinese EFL learners' oral language skills and to explore how learners' proficiency levels influence their learning outcomes.</p>
<p>Integrating TikTok in Higher Education: A Case of English as a Foreign Language Teacher Education Program (Benitez-Correa &amp; Gonzalez-Torres, 2025)</p>	<p>The study aimed to evaluate the effectiveness of TikTok videos in enhancing research and linguistic competencies among Latin-American EFL teacher education students and to explore their perceptions of TikTok as an educational tool.</p>

The article entitled “Second Language Task Engagement in Face-To-Face and Synchronous Video-Based Computer-Mediated Communication Modes: Performances and Perceptions” by Aubrey and Philpott (2023) investigates differences in EFL learners’ engagement across behavioral, cognitive, social, and emotional dimensions when performing decision-making tasks in synchronous video-based computer-mediated communication (SvCMC) and face-to-face (FTF) modes. The findings of this study indicate that face-to-face (FTF) mode significantly facilitated higher levels of behavioral, cognitive, social, and emotional engagement compared to synchronous video-based computer-mediated communication (SvCMC) mode. Specifically, FTF interactions elicited more turns, greater negotiation of content, increased affiliative backchannels, and higher enjoyment, reflecting enhanced interactivity, attention to interlocutors’ contributions, and a supportive classroom atmosphere. In contrast, SvCMC engagement was constrained by the absence of non-verbal cues, the cognitive load imposed by integrating

multiple stimuli, and reduced social presence, which collectively limited learners’ behavioral and social involvement. Notably, no significant differences were observed between FTF and SvCMC modes in the number of words produced, language-related episodes, or anxiety, suggesting that task modality predominantly influenced interactional dynamics and affective responses rather than linguistic attention per se. Learners’ perceptions further highlighted six factors shaping engagement—non-verbal cues, verbal interaction, stimuli integration, logistics, classroom atmosphere, and perceived learning value—underscoring the complex interplay between contextual, cognitive, and emotional dimensions in L2 task performance. These results reinforce the critical role of communication modality in shaping holistic engagement, demonstrating that FTF contexts provide richer opportunities for interaction, social connectedness, and positive emotional arousal, while SvCMC contexts, though effective, may require strategic pedagogical interventions to mitigate attentional and social limitations.

Research conducted by Mutluoğlu and Balaman (2023) entitled “The Use of Video Enhanced Observation in Video-Mediated Post-Observation Conversations on Pre-Service EFL Teachers’ Online Practicum Teaching” explores how video-mediated post-observation conversations, supported by a digital observation and video tagging tool, facilitate reflective practice and professional development for pre-service EFL teachers in online practicum teaching contexts. The findings of this study reveal that the integration of the Video Enhanced Observation (VEO) tool in a fully online, video-mediated post-observation meeting effectively facilitated pre-service teacher learning by creating a temporally and spatially bridged reflective environment that connected pre-teaching, while-teaching, and post-teaching phases. The tool enabled the teacher trainer to collaboratively review tagged teaching moments, particularly those involving the use of Kahoot! in EFL instruction, thereby extending reflection-on-action to reflection-for-action and reflection-in-action. Despite structural constraints imposed by the tightly organized feedback format, the pre-service teacher actively engaged in pedagogical

problem-solving, contingency planning, and self-assessment, demonstrating the affordances of VEO in supporting multimodal, evidence-based, and interactionally grounded reflective practices. Furthermore, the study highlights the critical role of mentor teachers in shaping pedagogical decisions before and during practicum teaching, emphasizing the co-construction of teacher learning outcomes through interactions among trainers, mentors, students, and digital tools. Overall, these results underscore that strategically implemented digital observation and video tagging tools can maximize teacher learning opportunities in fully online EFL settings, while necessitating a critical approach by all stakeholders to ensure alignment with institutional objectives and to optimize interactional space for trainee reflection beyond mere feedback delivery.

Research implemented by Alqefari (2024) entitled “The Presence of Translanguaging And Its Effect on EFL Learners’ Linguaging Opportunities in Video-Based Peer Editing” examines how translanguaging occurs in video-mediated peer editing among Saudi EFL learners and its impact on the production and resolution of language-related episodes (LREs). The findings of the present study indicate that translanguaging played a pivotal role in EFL dyads’ peer-editing activities, eliciting a greater number of language-related episodes (LREs) compared to English-only and Arabic-only interactional mediums, with the highest frequency observed in translanguaging LREs, followed by Arabic-only and English-only LREs. Analysis revealed that translanguaging facilitated the production of more elaborate LREs, particularly in addressing lexical and grammatical issues, while mechanical issues were predominantly resolved in L1, highlighting its scaffolding function in learners’ verification and negotiation of linguistic forms. Despite non-significant differences across peer interaction mediums in LRE focus, translanguaging enabled more successful resolution of LREs, as evidenced by the lower proportion of unresolved and unsuccessfully resolved LREs, underscoring its efficacy in promoting modified output and sustained engagement. Furthermore, the study corroborates sociocultural and interactionist perspectives, demonstrating that peer interaction—mediated through

translanguaging—supports cognitive development, attention to linguistic forms, and collaborative problem-solving. The integration of video-mediated peer editing provided a rich data environment for capturing dyadic dialogue, offering valuable pedagogical insights for EFL instruction by highlighting the importance of allowing learners autonomy in language choice during collaborative writing tasks. Overall, these results reinforce translanguaging as a strategic mechanism that enhances EFL learners’ awareness, negotiation, and resolution of linguistic issues, while emphasizing the instructional potential of digitally mediated peer interaction for form-focused language development.

Research carried out by Van, et al., (2024) entitled “Effects of Video-based Shadowing on Suprasegmental Features: EFL Learners’ Pronunciation Performance and Attitudes” investigates the effects of video-based shadowing on pre-intermediate EFL learners’ pronunciation of suprasegmental features and to explore their attitudes toward the technique. The present study demonstrates that video-based shadowing significantly enhances pre-intermediate EFL learners’ pronunciation of suprasegmental features, particularly intonation, linking, stress patterns, and pitch, with improvements more pronounced in controlled read-aloud tasks than in free-response activities. Post-intervention assessments revealed notable gains from initially low pretest scores, confirming that the technique effectively scaffolds learners’ phonological awareness and auditory-motor integration. Learners’ positive attitudes toward video-based shadowing were evidenced through questionnaires and interviews, highlighting perceived benefits such as improved accentedness, fluency, vocabulary expansion, and confidence in pronunciation, while challenges such as speaker rate, unfamiliar topics, and unknown vocabulary were identified. Pedagogically, these findings suggest that video-based shadowing should be systematically integrated into EFL instruction, particularly for learners with limited English proficiency or difficulties in producing suprasegmental features, with educators facilitating interactive pair or group activities, guiding video selection, and addressing common implementation challenges. The study underscores the potential of video-based shadowing as a targeted, evidence-based

approach for enhancing learners' pronunciation performance, advocating for its incorporation into curricula to support sustained, effective pronunciation development in EFL contexts.

Research undertaken by Wang and Lee (2025) entitled "The Impact of Video Dubbing App on Chinese College Students' Oral Language Skills Across Different Proficiency Levels" examines the effectiveness of an English video dubbing app in improving Chinese EFL learners' oral language skills and to explore how learners' proficiency levels influence their learning outcomes. The present study demonstrates that dubbing activities using a specialized app effectively enhance Chinese EFL students' oral language skills (OLSs), encompassing pronunciation, intonation, fluency, vocabulary, and grammatical accuracy, while simultaneously fostering learner engagement, autonomy, and confidence in speaking. Quantitative and qualitative analyses revealed that these tasks not only improved students' listening and speaking performance but also promoted proceduralization of articulatory knowledge through repeated practice and audiovisual synchronization, enabling learners to transform declarative knowledge into usable speaking skills. The integration of video, voiceover, and subtitles provided multimodal input that facilitated comprehension, vocabulary acquisition, and alignment of speech with visual cues, thereby addressing pronunciation fossilization and enhancing intonation, stress patterns, and pitch. Furthermore, the study found differential effects across proficiency levels, with lower-level learners demonstrating greater gains in pronunciation and higher-level learners in fluency and grammar-lexical use, highlighting the importance of differentiated task design. Pedagogically, the findings underscore the value of incorporating dubbing tasks into EFL curricula to create interactive, communicative, and student-centered learning environments, while emphasizing the need for tailored instructional strategies that accommodate varying learner proficiency, scaffold skill development, and sustain engagement. Overall, the study provides compelling evidence that app-based dubbing activities constitute an effective, multimodal, and adaptive approach to bridging the gap between traditional exam-oriented instruction

and communicative language practice, supporting meaningful and autonomous oral language development in EFL contexts.

Research conducted by Benitez-Correa and Gonzalez-Torres (2025) entitled "Integrating TikTok in Higher Education: A Case of English as a Foreign Language Teacher Education Program" evaluates the effectiveness of TikTok videos in enhancing research and linguistic competencies among Latin-American EFL teacher education students and to explore their perceptions of TikTok as an educational tool. The present study demonstrates that integrating TikTok videos into higher education courses significantly enhances students' academic competencies, as evidenced in both Contrastive Grammar and Research Methods in ELT courses. Quantitative analyses revealed that experimental groups exposed to TikTok interventions exhibited statistically significant improvements in performance compared to control groups, highlighting the platform's potential to provide engaging, multimodal learning experiences that facilitate retention and comprehension. Students' perceptions, captured through questionnaires and interviews, corroborated these findings, indicating that TikTok supports multimodal learning, promotes active and collaborative engagement, and fosters communication skills, critical thinking, and creativity through group-based video creation activities. While generally well-received, students noted challenges including potential distractions, comfort with audio-visual content, and technological limitations such as inconsistent internet access, emphasizing the need for careful integration and scaffolding to maximize effectiveness. Additionally, students recommended structured pre-lesson exercises, professionally edited and concise content, and increased involvement in content creation to enhance learning outcomes. Collectively, these results underscore that TikTok serves as a valuable supplementary educational tool, whose efficacy is optimized when combined with structured guidance, blended pedagogical strategies, and alignment with course objectives, ultimately facilitating active, student-centered, and multimodal learning in higher education contexts.

## DISCUSSION

### **Pedagogical Applications**

The pedagogical implications derived from recent studies highlight the transformative potential of multimodal and digitally mediated approaches in EFL instruction. The comparative analysis of task engagement across modalities underscores the pedagogical necessity of optimizing interactional spaces by leveraging the affordances of face-to-face contexts while strategically mitigating the attentional and social limitations of synchronous video-based communication. Likewise, the integration of video-enhanced observation exemplifies how digital tagging tools can foster reflective practice, enabling pre-service teachers to extend reflection-on-action toward reflection-in-action through evidence-based dialogue with mentors. Furthermore, the exploration of translanguaging in peer-editing activities provides compelling evidence for allowing strategic flexibility in language choice, thereby legitimizing L1 use as a scaffold for metalinguistic awareness, collaborative problem-solving, and successful resolution of linguistic issues. In addition, the use of video-based shadowing and app-based dubbing demonstrates the efficacy of audiovisual technologies in cultivating suprasegmental pronunciation features and oral language proficiency across varying levels of learners, emphasizing the importance of differentiated task design that addresses both lower-level phonological development and higher-level fluency and grammatical accuracy. Finally, the incorporation of short-form video platforms into teacher education programs illustrates how multimodal, student-generated content can significantly enhance linguistic competencies, creativity, and critical thinking when embedded within structured pedagogical frameworks. Collectively, these findings reaffirm that the effective integration of digital tools in EFL contexts requires not only technological innovation but also careful teacher mediation, differentiated scaffolding, and the cultivation of supportive social and emotional environments to ensure that learners' engagement, reflection, and communicative competence are holistically developed.

### **Emerging Developments**

Recent advancements in technology-enhanced language learning point toward

several emerging developments that are reshaping the landscape of EFL pedagogy. First, the convergence of multimodal learning platforms with artificial intelligence and adaptive technologies suggests a paradigm shift toward more personalized and data-driven instruction, wherein learner engagement, linguistic output, and affective responses can be dynamically monitored and scaffolded in real time. Second, the integration of video-mediated practices, such as observation, peer editing, shadowing, and dubbing, highlights a broader pedagogical movement toward authentic, situated, and interactionally rich learning experiences that bridge traditional classroom boundaries with digital environments. Third, the growing recognition of translanguaging practices signals a critical turn in language education, emphasizing the legitimacy of learners' full linguistic repertoires as valuable resources for collaborative knowledge construction and problem-solving. Moreover, the incorporation of popular digital platforms such as TikTok into higher education curricula represents a democratization of content creation and an expansion of student agency, creativity, and multimodal literacy, while simultaneously raising questions about digital ethics, distraction management, and equitable access. Looking forward, the trajectory of research and practice is likely to emphasize the fusion of AI-powered writing and speaking tools, immersive virtual and augmented reality environments, and learner-generated multimodal content as integral components of EFL learning ecologies. Collectively, these developments underscore the necessity for educators, researchers, and policymakers to adopt a critical yet innovative stance, ensuring that the affordances of emerging technologies are harnessed not merely for novelty but for sustainable, equitable, and pedagogically sound language learning outcomes.

### **Empirical Findings**

The corpus of studies under review provides a coherent body of empirical substantiation that illuminates the pedagogical affordances and constraints of digitally mediated, multimodal approaches to EFL learning and teaching. Aubrey and Philpott's (2023) comparative investigation of face-to-face vis-à-vis synchronous video-based communication demonstrates, through both



quantitative and qualitative analyses, that communicative modality exerts a profound influence on learner engagement across behavioral, cognitive, social, and emotional domains. Their findings reveal that while face-to-face contexts elicit richer interactional density, affiliative cues, and affective resonance, synchronous video-mediated modes are impeded by attenuated social presence and heightened cognitive load, thereby underscoring the necessity of pedagogical intervention to compensate for modality-induced limitations.

Parallel evidence is furnished by Kanat Mutluoğlu and Balaman (2023), whose study of Video Enhanced Observation (VEO) as a digital reflective scaffold illustrates how temporally and spatially distributed video-mediated interactions extend pre-service teachers' reflective practices beyond conventional feedback sessions. The empirical data highlight the role of VEO in facilitating reflection-in-action, enabling participants to reconstruct pedagogical decision-making through multimodal engagement with digitally tagged teaching episodes. Complementarily, Alqefari (2024) provides empirical validation for the pedagogical efficacy of translanguaging practices within video-mediated peer editing, demonstrating that bilingual interactional repertoires generate not only a higher incidence of language-related episodes but also a greater proportion of successfully resolved linguistic problems, thereby reinforcing the mediational value of L1 as a cognitive and collaborative resource in second language development.

The empirical rigor of interventionist approaches is further exemplified in Van et al. (2024) and Wang and Lee (2025), whose studies document the impact of audiovisual techniques on oral proficiency. Van et al. (2024) establish that video-based shadowing significantly enhances suprasegmental pronunciation features—intonation, stress, linking, and pitch—while simultaneously cultivating learner confidence, with effects more salient in controlled production tasks than in spontaneous speech. In a complementary fashion, Wang and Lee (2025) substantiate the efficacy of app-based video dubbing in advancing oral language performance across proficiency levels, empirically demonstrating differentiated gains wherein lower-level learners achieve pronounced improvements in

phonological accuracy while higher-level learners exhibit enhanced fluency and grammatical complexity.

Finally, Benitez-Correa and Gonzalez-Torres (2025) furnish both quantitative and perceptual evidence for the academic utility of TikTok integration in EFL teacher education programs. Their results reveal that engagement with student-generated multimodal content yields statistically significant improvements in research literacy, linguistic competencies, and critical-creative capacities, albeit tempered by challenges pertaining to distraction, technological constraints, and the necessity of pedagogical structuring.

Collectively, these empirical investigations converge to substantiate the claim that digitally mediated and multimodally enriched pedagogies constitute not merely ancillary tools but transformative mechanisms capable of enhancing linguistic competence, interactional engagement, reflective capacity, and creative expression. More critically, the empirical evidence underscores that such affordances are contingent upon context-sensitive design, differentiated scaffolding, and the active mediation of educators who can translate technological innovation into sustainable, equitable, and pedagogically robust learning outcomes.

### Research Gaps

Although the reviewed empirical studies provide compelling evidence for the pedagogical affordances of multimodal, digitally mediated, and app-based interventions in EFL contexts, several critical gaps remain that delimit the generalizability and sustainability of these findings. First, the majority of existing investigations, such as those by Aubrey and Philpott (2023) and Kanat Mutluoğlu and Balaman (2023), focus primarily on immediate engagement and reflective practices within controlled or short-term interventions. What remains underexplored is the longitudinal trajectory of such interventions, particularly whether the observed cognitive, affective, and interactional gains are sustained over extended periods and across diverse learning ecologies. Second, while studies like Alqefari (2024) and Van et al. (2024) highlight the efficacy of translanguaging and shadowing in specific linguistic domains, they tend to isolate discrete

competencies (e.g., pronunciation, LRE resolution) without sufficiently addressing how these skills interact holistically to influence broader communicative competence and academic literacy development.

Third, the geographical and demographic scope of the reviewed studies remains relatively narrow, with data predominantly drawn from single institutions or homogeneous learner populations (e.g., Saudi learners, Chinese college students, Latin-American teacher candidates). Consequently, there is limited understanding of how variables such as cultural background, educational policy, institutional affordances, and learners' socio-economic status mediate the effectiveness of digital and multimodal interventions. Fourth, despite emerging explorations of platforms such as TikTok (Benitez-Correa & Gonzalez-Torres, 2025) and dubbing applications (Wang & Lee, 2025), the field has yet to comprehensively examine issues of digital ethics, learner autonomy, and critical digital literacy, which are increasingly salient as learners engage with algorithm-driven, commercialized platforms. Finally, while current research convincingly demonstrates the pedagogical utility of digital tools, few studies adopt a comparative lens that directly investigates differential impacts across learners of varying proficiency levels, particularly with regard to how AI-powered tools or multimodal tasks may scaffold or constrain high- versus low-proficiency learners in achieving equitable outcomes.

Taken together, these gaps underscore the need for research that moves beyond short-term, skill-specific, and context-bound inquiries toward more integrative, longitudinal, and critically framed investigations. Future scholarship must interrogate not only the efficacy of emerging technologies in enhancing language skills, but also their role in shaping learner identity, agency, and critical awareness within increasingly globalized, digitally saturated educational landscapes.

### **Theoretical and Practical Implications**

From a theoretical perspective, the synthesis of these studies advances our understanding of how multimodal, digitally mediated, and AI-augmented practices reconfigure the processes of second language acquisition and teacher education. The

reviewed evidence underscores the necessity of reconceptualizing traditional frameworks of task engagement, reflective practice, and communicative competence through the lens of socio-cognitive, sociocultural, and multimodality theories. For instance, Aubrey and Philpott's (2023) comparative findings on FTF and SvCMC modes provide empirical substantiation for expanding task engagement theory to account for modality-specific cognitive load and affective dynamics. Similarly, Alqefari's (2024) findings on translanguaging validate the sociocultural perspective that learners' entire linguistic repertoires should be regarded as cognitive tools for mediating meaning-making and collaborative problem-solving. Moreover, the incorporation of video-mediated reflective tools (Kanat Mutluoğlu & Balaman, 2023) enriches teacher learning theories by demonstrating how reflection-on-action and reflection-in-action can be extended across temporally and spatially distributed learning environments, thereby broadening our understanding of teacher cognition in digital contexts. Collectively, these findings necessitate a theoretical shift toward recognizing language learning as a dynamic, multimodal, and technology-embedded practice that transcends conventional classroom boundaries.

On a practical level, the implications are equally significant for curriculum design, instructional strategies, and teacher professional development. The demonstrated efficacy of video-based shadowing (Van et al., 2024) and dubbing applications (Wang & Lee, 2025) suggests that audiovisual tasks should be systematically embedded into curricula to scaffold learners' oral proficiency, with differentiated scaffolding tailored to distinct proficiency levels. The evidence from translanguaging in peer editing further supports the pedagogical legitimacy of integrating L1 resources into collaborative tasks, especially in digitally mediated environments, as a means of enhancing problem-solving and linguistic accuracy. Additionally, the use of TikTok and other learner-generated multimodal platforms (Benitez-Correa & Gonzalez-Torres, 2025) points to the practical necessity of rethinking classroom participation, creativity, and critical literacy in ways that align with learners' digital lifeworlds, while simultaneously requiring

educators to address issues of digital ethics, distraction management, and equitable access.

For teacher education, the integration of Video Enhanced Observation (Kanat Mutluoğlu & Balaman, 2023) demonstrates that digital tagging and feedback tools can significantly enrich reflective practice by providing evidence-based, multimodal records of teaching performance. This suggests that teacher training programs should prioritize the development of digital pedagogical literacy, equipping future teachers with the capacity to critically engage with, adapt, and repurpose technological tools in ways that foster learner agency and optimize learning outcomes. At the policy level, the findings call for institutional frameworks that support sustainable technological integration by providing infrastructure, professional training, and context-sensitive guidelines to ensure equitable access and pedagogical coherence.

In sum, the theoretical and practical implications of this body of research converge to highlight a transformative paradigm in EFL education—one in which multimodal, digitally mediated, and AI-supported practices are not ancillary supplements but central mechanisms that reconfigure how learners engage with language, how teachers conceptualize reflection and scaffolding, and how institutions design equitable and sustainable learning ecologies.

### **Future Research Directions**

Building on the identified gaps and the theoretical–practical insights synthesized from the reviewed studies, several directions for future inquiry emerge that warrant sustained scholarly attention. First, longitudinal investigations are urgently needed to determine the durability and transferability of learning gains achieved through multimodal, digitally mediated interventions. While existing studies (e.g., Aubrey & Philpott, 2023; Kanat Mutluoğlu & Balaman, 2023) highlight immediate effects on engagement and reflective practices, little is known about whether these cognitive, affective, and interactional benefits persist over time, transfer to broader communicative competence, or extend to diverse sociocultural and institutional contexts.

Second, future research should adopt a more integrative perspective that moves

beyond discrete-skill foci—such as pronunciation (Van et al., 2024) or peer-editing LREs (Alqefari, 2024)—to examine how multimodal interventions shape holistic language development. This entails exploring the synergistic interplay among oral proficiency, academic literacy, metacognitive regulation, and critical digital literacy, particularly as learners navigate complex digital ecosystems that blur the boundaries between formal and informal learning spaces.

Third, there is a pressing need to broaden the demographic and geographical scope of inquiry. The predominance of context-specific samples, such as Saudi, Chinese, or Latin-American EFL learners, restricts the generalizability of findings. Cross-cultural, cross-institutional, and comparative studies could illuminate how sociocultural backgrounds, educational policies, and institutional affordances mediate the effectiveness of digital and multimodal pedagogies, while also shedding light on issues of equity, inclusivity, and learner identity.

Fourth, future scholarship should critically interrogate the role of emerging technologies—particularly AI-powered tools, immersive VR/AR environments, and algorithm-driven social media platforms—not only in enhancing language performance but also in shaping learner agency, autonomy, and ethical awareness. With the increasing reliance on commercialized and data-driven platforms such as TikTok, the field must address questions of digital ethics, surveillance, and the commodification of learner data, while simultaneously developing pedagogical frameworks that empower learners to engage with such tools critically and responsibly.

Finally, there is a need for nuanced, proficiency-sensitive research designs that examine how learners at varying levels of linguistic competence engage with and benefit from technology-enhanced interventions. As evidenced by Wang and Lee (2025), learners of different proficiency levels respond differentially to multimodal tasks, suggesting that future studies should adopt stratified or mixed-method designs to capture these variations more systematically. Such work would provide empirically grounded insights for tailoring instructional strategies and scaffolding mechanisms that ensure equitable learning outcomes across proficiency spectra.

Taken together, these future directions call for a shift toward longitudinal, integrative, cross-contextual, ethically informed, and proficiency-sensitive research paradigms. Such scholarship will not only extend the empirical robustness of current findings but also contribute to the development of sustainable, equitable, and critically aware digital pedagogies in EFL education.

## Conclusion

The synthesis of recent empirical investigations unequivocally demonstrates that the integration of multimodal, digitally mediated, and app-based interventions constitutes a transformative paradigm in EFL pedagogy. Across diverse contexts, these approaches have been shown to significantly enhance learners' engagement, reflective capacities, pronunciation, oral proficiency, and critical literacy, while simultaneously fostering creativity, autonomy, and professional development. Collectively, the findings affirm that technological affordances—ranging from synchronous video-based communication and video tagging tools to translanguaging in peer editing, video shadowing, dubbing applications, and social media platforms such as TikTok—are not peripheral supplements to traditional instruction but central mechanisms for cultivating holistic communicative competence.

However, the reviewed body of work also underscores the conditional nature of these affordances: their efficacy is inextricably linked to pedagogical mediation, differentiated scaffolding, and context-sensitive design. Without such mediation, digital tools risk amplifying cognitive load, attenuating social presence, or exacerbating inequities in access and participation. This recognition situates the role of the teacher not merely as a facilitator of technological adoption but as a critical mediator who aligns innovation

with learners' needs, institutional goals, and ethical imperatives.

At the theoretical level, the convergence of socio-cognitive, sociocultural, and multimodality perspectives invites a reconceptualization of language learning as a dynamic, technology-embedded process that transcends conventional classroom boundaries. Practically, the evidence mandates a systematic integration of multimodal and AI-supported practices into curricula, professional development programs, and institutional policies that prioritize sustainability, equity, and critical digital literacy.

In light of the identified research gaps, the trajectory of future inquiry must shift toward longitudinal, cross-contextual, and proficiency-sensitive paradigms that investigate not only skill-specific outcomes but also the broader development of learner identity, agency, and critical awareness within digitally saturated ecologies. Such scholarship will be instrumental in ensuring that the promises of emerging technologies are harnessed responsibly and sustainably, yielding equitable and pedagogically robust outcomes for diverse EFL populations.

Ultimately, this body of research converges on a central conclusion: the future of EFL education lies in strategically leveraging digital and multimodal affordances to cultivate learners' communicative competence, reflective capacities, and critical literacies, while simultaneously safeguarding inclusivity, sustainability, and ethical engagement in an increasingly globalized and technologically mediated world.

## References

Alqefari, A. N. (2024). The presence of translanguaging and its effect on EFL learners' languaging opportunities in video-based peer editing. *System*, 125.

<https://doi.org/10.1016/j.system.2024.103452>

Aubrey, S., & Philpott, A. (2023). Second language task engagement in face-to-face and synchronous video-based computer-mediated communication modes: Performances and perceptions. *System*, 115. <https://doi.org/10.1016/j.system.2023.103069>

Avsheniuk, N., Lutsenko, O., Seminikhyna, N., & Svyrydiuk, T. (2023). Fostering Intercultural Communicative Competence and Student Autonomy through Project-Based Learning. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.4348413>

Benitez-Correa, C., & Gonzalez-Torres, P. (2025). Integrating TikTok in Higher Education: A Case of English as a Foreign Language Teacher Education Program. *Educational Process: International Journal*, 15. <https://doi.org/10.22521/edupij.2025.15.135>

Brewin, J. A. (2025). *The Aural Stage: Radio Drama as Sonic Lazzi*. <https://doi.org/10.11575/PRISM/48901>

Choi, J., & Nunan, D. (2022). Learner contributions to materials in language teaching. *The Routledge Handbook of Materials Development for Language Teaching*, 429–440. <https://doi.org/10.4324/B22783-37/LEARNER-CONTRIBUTIONS-MATERIALS-LANGUAGE-TEACHING-JULIE-CHOI-DAVID-NUNAN>

Duque, R. de C. S., Monteiro, R. R., Loureiro, V. J. S., Nascimento, I. J. B. M. F. do, Placido, R. L., Silva, C. J. da, & Saraiva, M. do S. G. (2024). Training teachers to use technology: Artificial Intelligence (AI) and the new challenges facing education. *Themes Focused on Interdisciplinarity and Sustainable Development Worldwide V. 02*. <https://doi.org/10.56238/SEVENED2024.003-019>

Grandon, M. (2018). *Exploring the use of video-based materials in the Japanese university English language classroom*.

Jain, P., Maithilisharan, R., Mahavidyalaya, G., & Chirgaon, J. (2024). Authentic Materials for Intercultural Communication: Develop Speaking Materials That

Encourage Learners to Engage in Authentic Cross-Cultural Communication Effectively in Diverse Global Contexts. *Creative Saplings*, 3(4), 1–14. <https://doi.org/10.56062/gtrs.2024.3.4.563>

Kanat Mutluoğlu, A., & Balaman, U. (2023). The use of Video Enhanced Observation in video-mediated post-observation conversations on pre-service EFL teachers' online practicum teaching. *System*, 118. <https://doi.org/10.1016/j.system.2023.103151>

Khalima, D., & Teacher, S. (2025). THE ROLE OF VIDEO-BASED INSTRUCTION IN FOREIGN LANGUAGE ACQUISITION: PEDAGOGICAL IMPLICATIONS AND EFFECTIVENESS. *Web of Teachers: Inderscience Research*, 3(4), 236–240. <https://webofjournals.com/index.php/1/article/view/3908>

Li, Z., & Zhang, W. (2025). Technology in education: Addressing legal and governance challenges in the digital era. *Education and Information Technologies*, 30(7), 8413–8443. <https://doi.org/10.1007/S10639-024-13036-9/METRICS>

Nelson, B. C., & Erlandson, B. E. (2008). Managing cognitive load in educational multi-user virtual environments: Reflection on design practice. *Educational Technology Research and Development*, 56(5–6), 619–641. <https://doi.org/10.1007/S11423-007-9082-1/METRICS>

Nowak, S., & Kaplan-Rakowski, R. (2024). The Impact of Relaxational Virtual Reality and Video Learning Activities on Subjective Vitality in a Blended Learning Environment. *International Journal on E-Learning: Corporate, Government, Healthcare, and Higher Education*, 23(4), 415–428. <https://doi.org/10.70725/377570TSZDGT>

Oyelana, O. O., Olson, J., & Caine, V. (2022). An evolutionary concept analysis of learner-centered teaching. *Nurse Education Today*, 108, 105187. <https://doi.org/10.1016/J.NEDT.2021.105187>

Palaigeorgiou, G., & Papadopoulou, A. (2019). Promoting self-paced learning in the

elementary classroom with interactive video, an online course platform and tablets. *Education and Information Technologies*, 24(1), 805–823. <https://doi.org/10.1007/S10639-018-9804-5/METRICS>

Sain, Y., & Hermansyah, S. (2025). Exploring the Language Attitudes of the Tolaki Community in Kendari: A Comprehensive Sociolinguistic Analysis. *Journal of Languages and Language Teaching*, 13(2), 983–993.

<https://doi.org/10.33394/jollt.v13i2.14415>

Sam Hermansyah, Buhari, Ibrahim Manda, Andi Sadapotto, Muhammad Hanafi, Andi Asrifan, ... Muliani. (2024). Reflection on Learning in Teacher Professional Education (PPG) Strategies to Enhance the Competence of Novice Teachers at Universitas Muhammadiyah Sidenreng Rappang. *INTERACTION: Jurnal Pendidikan Bahasa*, 11(2), 727–734.

<https://doi.org/10.36232/interactionjournal.v11i2.628>

Smyth, R. (2011). Enhancing learner–learner interaction using video communications in higher education: Implications from theorising about a new model. *British Journal of Educational Technology*, 42(1), 113–127. <https://doi.org/10.1111/J.1467-8535.2009.00990.X>

Van, P. T. T., Lien, L. K., & Khang, N. D. (2024). Effects of Video-based Shadowing on Suprasegmental Features: EFL Learners' Pronunciation Performance and Attitudes. *REFlections*, 31(3), 896–924. <https://doi.org/10.61508/refl.v31i3.275498>

Wang, X., & Lee, S. M. (2025). The impact of video dubbing app on chinese college students' oral language skills across different proficiency levels. *International Journal of Educational Research*, 130. <https://doi.org/10.1016/j.ijer.2024.102521>