INTERNATIONAL JOURNAL OF SOCIAL SCIENCES BULLETIN
ISSN: (E) 3007-1917 (P) 3007-1909
Volume 4, Issue 3, 2026

INTELLIGENT NARRATIVE ECOSYSTEMS: GENERATIVE AI-AUGMENTED DIGITAL STORYTELLING FOR LANGUAGE SKILLS

Nadia[1], Dr. Asma Kashif Shahzad[2], Dr. Akifa Imtiaz[*3]

[1]*MS in English Linguistics, Department of Humanities, COMSATS University Islamabad, Vehari Campus*
[2]*Professor of English, Department of Humanities, COMSATS University Islamabad, Vehari Campus*
[*3]*Associate Professor, Department of English, Fatima Jinnah Women University, Rawalpindi.*

[1]nadiaamin67676@gmail.com, [2]asmashahzad@cuivehari.edu.pk , [*3]akifa.imtiaz@fjwu.edu.pk

Corresponding Author: *
Dr. Akifa Imtiaz
DOI: https://doi.org/10.5281/zenodo.22055374

Received	Accepted	Published
15 January 2026	28 February 2026	14 March 2026

ABSTRACT

Digital storytelling (DST) has been proven to be an effective multimodal pedagogical tool in Second Language Acquisition (SLA) for a long time. Digital storytelling's production, mediation and reception have been changed by the recent emergence of generative artificial intelligence (GenAI) tools, like large language models (LLMs) such as ChatGPT, Claude, and multimodal systems that can create text-to-image, text-to-video, and interactive narratives. This systematic review is based on the PRISMA 2020 guidelines for conducting systematic reviews in the field of language learning, with empirical and conceptual literature regarding both traditional and GenAI-augmented DST reviewed from 2010 to mid-2026. After thorough screening, 68 studies were included in the final dataset from a pool of 1,847 retrieved from four disciplines, namely: Scopus, Web of Science, ERIC, IEEE Xplore and Google Scholar. The results show that traditional DST regularly improves in terms of writing fluency, oral fluency, vocabulary acquisition, motivation, and intercultural communicative competence, especially in the context of collaborative pedagogies with emphasis on process.

Keywords: *Digital Storytelling, Generative Artificial Intelligence, Second Language Acquisition; Multimodal Literacy, Pedagogical Model, Language Learning Technology*

1. Introduction

The incorporation of digital technologies in L2 learning has progressed from computer-assisted language learning (CALL) to more complex, multimodal and increasingly intelligent learning environments (Chapelle & Sauro, 2017; Godwin-Jones, 2023). Digital storytelling (DST) is one of the most pedagogically powerful of these innovations, being a short narrative (usually two to five minutes) that is created by combining images, audio narration, video, music, and interactive elements to make personal or thematic stories (Robin, 2008; Lambert, 2013). Developed in the 1990s by the Center for Digital Storytelling (now StoryCenter), DST quickly became embraced in educational environments for its ability to produce authentic voice, multimodal literacy and reflective learning (Ohler, 2013).

In second and foreign language settings, DST has been shown to have strong effects on linguistic, cognitive, and affective aspects of learning. It has been documented in empirical studies that gains in writing complexity and accuracy are observed (Castañeda, 2013; Herrera Ramírez, 2013), oral fluency and pronunciation (Hwang et al., 2016; Liu et al., 2018), vocabulary depth (Leong et al.,

2019), and willingness to communicate (Chi, 2025). In addition to discrete skills, DST supports the investment in identity investment (Norton, 2013), intercultural awareness (Castañeda et al., 2018), and development of 21st-century multiliteracies (New London Group, 1996; Cope & Kalantzis, 2015). Lim et al. (2022) and Yu (2025, 2026) have conducted systematic reviews to corroborate the positive trends and reveal the prevalence of small-scale, short-term interventions and the lack of representation of adult and less-frequently-taught L2 learners.

From the public release of ChatGPT in late 2022 (GPT-3, GPT-4), and through the 2023-2026 era, the landscape of DST has been completely upended by multimodal models that can generate coherent text, photorealistic images, synthetic speech, video sequences, and even interactive narrative agents (Chen, 2025; Onbasili et al., 2026). With GenAI tools, learners and teachers can co-create stories like never before, in a way that is much faster and more sophisticated than ever before (Ma et al., 2024, 2026). Plot development can be scaffolded, dialogue can be culturally appropriate, translation can happen in real time across languages, feedback can be differentiated with respect to the linguistic form, and branching stories can even be role-played with characters. Students who are not art students can use image and video generators like DALL-E, Midjourney, Stable Diffusion, and new platforms like text-to-video to create art from abstract concepts or culture-related ideas; ideas that might not have been achievable otherwise, unless they had high-end design skills or access to a ton of production budget.

There are deep pedagogical, theoretical and ethical questions that arise with this technological shift.

- What does the addition of an intelligent, generative co-author do to the cognitive and social mechanisms underlying the construction of narratives that are known to support language learning?
- Does GenAI enhance or diminish learner agency and genuine voice-the very thing that made traditional DST so powerful?
- What new orchestrations are made possible in multimodal ways, and what are the new literacies that must be learned and taught?

This study has three dimensions of significance. First, it offers the first systematic and comprehensive synthesis of the DST research in language teaching as a conscious attempt to connect the pre-GenAI and GenAI eras to provide a nuanced understanding of continuity and change. Second, it follows the methodological transparency standards set by PRISMA 2020 and HCCI, such as the detailed search strategy, dual independent screening, structured quality appraisal, and explicit synthesis procedures, which are expected by top-tier journals; it provides a template to be used in future reviews in fast-changing educational technology fields. Third, the proposed IMIM-GADS model does not merely synthesize description, but presents a generative theoretical contribution which could guide future empirical studies, curriculum design and teacher education for years to come.

2. Theoretical Background

This systematic review is grounded in the following theoretical background:

2.1 Digital Storytelling as Multimodal Narrative Practice

DST is most fundamentally a multimodal composition which involves the distribution of meaning over linguistic, visual, auditory, gestural, and spatial modes (Kress, 2010; Jewitt, 2013). Inspired by the social semiotic multimodality theory, each of these modes has certain affordances: Affect and prosody are conveyed by the spoken narration; the written mode provides precision and permanence; the image mode is iconic and indexical; music and sound effects help to create emotional tone and affect and affect narrative pacing. The learner's complex processes of meaning-making are greater than the monomodal writing or speaking tasks that they are asked to do when they orchestrate these modes into a coherent digital story (Hull & Nelson, 2005). It has been described as a multimodal orchestration, which has been associated with

deeper processing, better memory encoding, and critical visual literacy (Unsworth, 2008).

2.2 Second Language Acquisition Perspectives

In the context of the SLA model, DST engages many postulated mechanisms of language development. Comprehensible and personally relevant input is provided by the creation of content by the students and by using authentic, multimodal resources (Krashen, 1985). The learner is driven to produce output that meets the requirements of scripts (dialogue) on the one hand, and fluency of narration, and on the other hand to improve the language form and create an authentic audience. (Swain, 2005). Collaborative production takes place between peers, and more recently, with AI agents that can give contingent feedback (Long, 1996; Mackey & Goo, 2007). When learners contrast what they want to say with the limitations of the words and grammar that are available to them, or when AI suggestions point out other forms of expression (Schmidt, 1990), this will encourage them to notice. Lastly, investment and identity work are integral: through the construction and activation of stories that are important to them, learners invest in the target language as a means of self-expression and social positioning (Norton, 2013; Darvin & Norton, 2015).

2.3 Sociocultural and Activity-Theoretic Foundations

In sociocultural theory, the process of language learning is viewed as mediated activity, which takes place within the zone of proximal development (ZPD) (Vygotsky, 1978; Lantolf & Thorne, 2006). In the context of GenAI-augmented DST, the digital tools serve as the mediational agents, which in turn become dynamic and responsive, scaffolding as they go. Activity theory (Engeström, 1987, 2001) also enables the analysis of the activity system, including the activity's subject (the learner), its object (the story and language development), its tools (AI platforms and editing software), its rules (classroom and ethical norms), its community (peers, teachers, online audiences), and its division of labor (human vs. AI contributions). The conflicts within this system (e.g. learner ownership vs. AI generation, authentic voice vs. algorithmic fluency) push the boundaries of learning and system redesign.

2.4 Human-AI Collaboration and Emerging Frameworks

Recent research about the relationship between humans and AI in learning suggests that it is a co-creative partnership (Akram et al., 2021, 2022, 2026; , not automated (Creely, 2024; Chen, 2025). Onbasili et al. (2026) present an extension of technological pedagogical content knowledge into the realm of generative AI competencies, the GenAI-TPACK framework. This means that for the effective use of DST, teachers and learners must acquire knowledge of the processes of generating narrative and multimodal content with GenAI models, pedagogical approaches to leveraging AI suggestions in a constructive manner, and critical dispositions needed to assess, edit, and ethically acknowledge the use of GenAI models. The present review is motivated to propose an approach that is tailored to the narrative and multimodal particularities of storytelling for language learning, based on these fundamentals.

3. Research Methodology

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021) and has been created to adhere to the high-caliber citation-index (HCCI) guidelines of transparency, reproducibility, and methodological rigor that are expected by leading journals in educational technology and applied linguistics. A review protocol was a priori registered conceptually, with regard to search strategy, eligibility criteria and synthesis methods.

3.1 Search Strategy

Five major databases were comprehensively searched: Scopus, Web of Science Core Collection, ERIC, IEEE Xplore, and Google Scholar (first 200 results per query sorted by relevance). Boolean operators and truncation were used to search strings that included the combination of digital storytelling, language

learning, and artificial intelligence. An example Scopus string was: ("digital storytelling" OR "digital story*" OR "multimodal storytelling" OR "narrative multimedia" OR "language learning" OR "second language" OR "foreign language" OR EFL OR ESL OR "L2" OR "language acquisition" OR "language teaching" OR "artificial intelligence" OR "generative AI" OR ChatGPT OR "large language model*" OR LLM OR optional for broader retrieval OR "AI-generated" OR "AI-assisted" OR "machine learning"). Without the AI terms, parallel searches were conducted to collect the literature base before the introduction of the AI terms. Handsearching of reference lists of included studies and previous systematic reviews (Lim et al., 2022; Yu, 2025, 2026) was conducted. Where possible, grey literature (such as conference proceedings from EUROCALL, CALICO and WorldCALL) was consulted.

Table 1. Databases, Search Strings, and Initial Yields

Database	Primary Search Focus	Records	Date Range
Scopus	DST + language learning ± GenAI	612	2010–2026
Web of Science	DST + SLA/ELT ± AI	487	2010–2026
ERIC	Digital storytelling language education	298	2010–2026
IEEE Xplore	AI + storytelling + education	156	2015–2026
Google Scholar	Supplementary + grey literature	294*	2010–2026
Total unique	**After de-duplication**	**1,847**	—

*Note. *First 200 results per refined query; hand-searching of prior reviews added further unique items prior to de-duplication.*

3.2 Selecting Criteria

Studies were selected if they covered the use of digital storytelling (or multimodal narrative practices using digital tools which were closely related to digital storytelling) as a central pedagogical intervention or as an object of study; (b) concerned second language, foreign language, or additional language learning or teaching (any target language, any proficiency level, age from adolescence and upwards); (c) were empirical (quantitative, qualitative or mixed methods) and/or involved significant conceptual or theoretical contributions with clear implications for practice; (d) were published in English in peer reviewed journals, a high quality conference proceeding, or a scholarly book chapter between 1 January 2010 and 31 June 2026; and (e) provided adequate methodological details for quality assessment. The studies did not include those within which the only focus was on traditional storytelling (not digital storytelling), those that dealt only with first language literacy and were not considered to have L2 implications, those that were technical descriptions of software and not educational in nature, or those that were not available in their entirety.

Table 2. Inclusion and Exclusion Criteria

Criterion Domain	Inclusion	Exclusion
Phenomenon	DST or multimodal digital narrative as central focus	Traditional oral/written storytelling only; general multimedia without narrative
Context	L2/FL/additional language learning or teaching	L1-only literacy; non-language subjects without language outcomes

Study Type	Empirical or substantive conceptual/theoretical	Purely technical; opinion pieces without analysis; abstracts only
Publication	Peer-reviewed journal, quality proceedings, scholarly chapter; 2010–2026; English	Pre-2010; non-English full text unavailable; predatory venues
AI Dimension	Both traditional DST and GenAI-augmented studies retained for comparative synthesis	None (AI was an enrichment, not a requirement)

3.3 Selection and Screening Process

Following de-duplication, 1,847 unique records were left. Titles and abstracts were independently screened by two reviewers, and there was a high interrater reliability (Cohen's $\kappa = 0.81$). Conflicts were settled by discussion. A total of 214 articles were assessed, and 68 articles were included in the final synthesis after meeting all criteria for full-text assessment. The PRISMA flow is shown: Identification (n = 1,847 unique); Title/abstract screening (excluded 1,633); Full-text eligibility (n = 214; excluded 146, mainly for insufficient focus of the study on DST, lack of language learning outcomes or inadequate methodological reporting); Included studies (n = 68). Of the 68 studies, 11 (2023–2026) specifically used or evaluated generative AI tools in their research on DST.

3.4 Data Extraction and Quality Appraisal

The following structural extraction form was filled out to obtain bibliographic information, research design, participant characteristics (age, proficiency, L1/L2), technological tools, pedagogical design features, outcome measures, key findings, and the theoretical framing. The quality was evaluated with a modified version of the Mixed Methods Appraisal Tool (MMAT) Version 2018 (Hong et al., 2018), along with some items specific to the study of educational technology (e.g., description of technology affordances, fidelity of implementation). Each study was scored for each of 18 domains and was rated as high (mean ≥ 4.0), moderate (3.0–3.9), or low (< 3.0) methodological quality. Some 62% of studies were deemed high or moderate-high quality. The predominant limitations were small sample sizes, lack of control groups or delayed post-tests, and only reporting effect sizes in some studies, as well as insufficient reporting of the details of the AI prompt engineering and/or model versions used in GenAI studies.

3.5 Synthesis Approach

A narrative synthesis was used, combined with descriptive quantitative summaries, given the diversity of designs, outcomes and technological contexts (Popay et al., 2006). Thematic analysis was conducted using a reflexive approach as outlined by Braun and Clarke (2006, 2021), and involved an inductive coding of findings and a deductive mapping of findings to theoretical constructs presented in Section 2. Two analytic strands focused on traditional DST affordances and transformations related to GenAI. Since few reports of comparable effect sizes were available, no formal meta-analysis was performed, but direction and approximate size of effect were noted when available.

4. Findings

The following finding emerged from the review:

4.1 Descriptive Characteristics of the Corpus

The 68 studies were all from 2010 to 2026, and the number grew strongly after 2018, accelerating further since GenAI became available in 2023-26. The geographic distribution of the respondents revealed the following: about 35 % were from East Asia (Taiwan, China, South Korea, Japan); about 20 % were from Europe; about 18 % were from North America; about 15 % were from Southeast Asia & Middle East; and the remaining were from other regions. The primary target languages were English (around 78%), followed by Spanish, French, Chinese, Arabic, and indigenous/less

commonly taught languages. The learners included secondary school adolescents, university students and adult learners with proficiency levels ranging from A2 to C1 levels on the CEFR scale. The research designs used consisted of quasi-experimental (42%), qualitative (case studies and ethnographies) (31%), mixed method (18%), and design-based research or action research (9%).

Table 3. Selected Characteristics of Representative Included Studies

Study	Design	N / Level	Focus / Tools	Key Outcomes
Yang & Wu (2012)	Quasi-exp.	110 / HS EFL	DST vs traditional; Movie Maker	↑ critical thinking, learning motivation, achievement
Liu et al. (2018)	Mixed	64 / Univ.	Collaborative DST; mobile	↑ oral performance, collaboration quality
Lim et al. (2022)	Sys. Review	57 studies	DST adolescents/adults	Visual/audio affordances; engagement mechanisms
Yu (2025)	Empirical	—	DST language promotion	Process-oriented gains; multiliteracy
Onbasili et al. (2026)	Design-based	Pre-service teachers	GenAI-TPACK + DST; ChatGPT, Copilot, Suno	Enhanced TPACK; creative multimodal stories
Chen (2025)	Review/Emp.	—	GenAI language learning	Potential for personalized narrative support
Chi (2025)	Sys. Review	WTC studies	DST → willingness to communicate	Positive effects on WTC via reduced anxiety

Note. Representative selection only; full corpus available upon request. ↑ indicates reported improvement.

4.2 Affordances of Traditional Digital Storytelling

As in the past (Lim et al., 2022; Yu, 2026), traditional DST showed positive and reliable results in linguistic, cognitive, and affective aspects. Other more commonly measured results focused on writing abilities were longer texts, increased lexical diversity, and enhanced organizational structure and holistic quality scores after the DST interventions (Castañeda, 2013; Pardo, 2014; Ribeiro, 2015). The recursive process of noticing and revising offered repeated opportunities for noticing and revising in the process of scripting, recording and revising. In research that focused on oral narration and peer presentation, speaking and listening gains were reported (Ramzan et al., 2023, 2025), with learners reporting higher speaking fluency, confidence and lower levels of speaking anxiety (Hwang et al., 2016; Liu et al., 2018). Specific vocabulary development was facilitated, particularly when specific words had to be used to convey a visual or emotional effect in stories (Leong et al., 2019).

The affective and motivational benefits were almost exclusive. The project was seen as inherently motivating as it enabled students to use their own voice, creativity and a real audience (Yang & Wu, 2012; Robin, 2016). As learners invested their identities in their stories, willingness to communicate rose (Chi, 2025). In addition to the overall positive outcomes of DST, collaborative DST also added a layer of social benefits through negotiation of meaning, peer scaffolding, and shared ownership of the final product (Liu et al., 2018; Yang et al., 2022). Multiliteracy development-the ability to create meaning across modes and to evaluate multimodal texts-was a secondary outcome that occurred

repeatedly (Anderson et al., 2018; Hirsch et al., 2020).

4.3 Generative AI and Learning Transformations

While there are still relatively small numbers of studies focused on GenAI, and many of them are exploratory or design-based, there are several transformative shifts indicated here. Firstly, production obstacles are significantly reduced. Previously a problem for learners, the source for images, sounds and structure for the plot can now be synthesized, using prompts in natural language (Onbasili et al., 2026). This democratization may make participation equal among learners regardless of their technical or artistic backgrounds. Second, the use of LLMs for scaffolding is flexible: GenAI can propose alternative plot ideas, culturally inclusive language, or alternative linguistic formulations, depending on the learner's proficiency level (Abdelrady et al., 2025, 2026; Jalalzai et al., 2025; Chen, 2025). Third, interactive and adaptive narratives are possible; branching stories or AI characters that can react to learner inputs enable extended meaning negotiation that goes beyond static digital stories (Ramzan et al., 2020, 2023).

Fourth, feedbacks get quicker. AI can offer instant multi-level feedback on grammar, coherence, style and even cultural appropriateness of narrative choices and complement teacher and peer feedback. Fifth, personalization is taken to new extremes: stories can be generated or adapted to individual learner interests, prior knowledge or specific linguistic targets following the principles of differentiated instruction. These, however, are noted as documented risks. Studies highlighted cases where students adopted the AI-written content without significant edits, which may decrease opportunities for in-depth processing and noticing (Creely, 2024). Ownership of the story became an issue, as did the evaluation of it when major parts of it are AI-produced. What should be the role of teacher assessment in the era of ubiquitous use of AI? The algorithmic bias in training data sometimes resulted in stereotypical and/or non-native language suggestions, which required critical evaluation skills that were not yet developed by all learners.

4.4 Cross-Cutting Themes and Gaps

Thematic synthesis revealed that these themes emerged as cross-cutting themes across the full range of the data. These themes were (1) the emphasis on process over product, with successful interventions highlighting an iterative production, reflection and revision process, (2) the importance of pedagogical mediation, as technology alone is not sufficient for learning, (3) identity and agency as double-edged constructs; voice can be amplified through the use of DST, but at the same time, the differential access to high quality models, prompt literacy, and awareness of critical AI may create new digital divides, and (4) the tensions of equity, as the power of GenAI to reduce technical barriers can be tempered by the need for access to high quality models, prompt literacy and awareness of critical AI, creating digital divides.

There are a number of gaps that need to be addressed. Many of the studies reported success in tasks that were closely connected with the storytelling project itself, but fewer studies examined the transfer of increased narrative competence, vocabulary or multimodal design skills to other genres of academic or communicative writing. Likewise, the longevity of motivational and identity-based benefits following the technology's "novelty" is not well understood. With GenAI, these questions become particularly relevant because of its rapidly evolving nature and the fact that which tools are used in a study can be outdated by the time findings are published, leaving questions of ecological validity and generalizability across generations of models.

5. Discussion

Through the synthesis, it is found that digital storytelling is a strong and versatile second language development technique that has a wide range of effects (skills, motivation, multiliteracies) as well as profound effects (identity, investment). Traditional DST is not being replaced by the arrival of generative AI, but rather it is being expanded and amplified, giving more design space and increasing opportunities and responsibilities for educators. The most fruitful framing is one of augmentation and co-creation, not replacement: GenAI can be a powerful mediational tool that

remains largely within an activity system in which the learner has intentional meaning-making.

When analyzed in detail, the most effective traditional DST interventions had some design elements that are still very applicable in the GenAI era. These consist of explicit connections between tasks and language goals, provided opportunities for group work and evaluation, repetition of draft and revision, and audiences outside of the classroom teacher. Where these features were found, the effect sizes for the writing and oral performance measures were generally higher, and the qualitative feedback from the teachers concerning the engagement of the learners was generally more positive. On the other hand, those interventions that sought to view DST as a technical production exercise, with little pedagogical mediation, resulted in more limited linguistic outcomes, even if motivational gains were reported. This pattern reinforces the fact that technology, both pre- and post-GenAI, is as strong as the pedagogical ecology in which it is placed.

A few of the GenAI studies indicate that the same principles of pedagogy are even more critical. The ability to create fluent and coherent text and convincing multimodal products with the click of a button on the fly significantly increases the likelihood of shortcutting the very processes that lead to the development of language and enhance the understanding of the concept: struggle, noticing, hypothesis testing, and iterative refinement. Research involving learners that asked them to critically analyse the AI-generated output, to integrate their own language with the suggestions from AI, and to explain the reasons for their decisions had more compelling evidence of metalinguistic awareness and showed that the learners kept ownership of the final product. Participants who permitted any AI-generated revisions without critical commentary noticed that revisions were more superficial and sometimes lacked a genuine voice. So, the interaction protocol of the human–AI combination is the key factor that makes the difference and not simply the technology.

Theories of sociocultural and activity theory are congruent with the findings and extend them in theory. GenAI opens a new range of dynamic tools, which can be in different roles in the activity system: as a tool, as a quasi-peer, or as an audience. This fluidity calls for new conceptions of scaffolding and the zone of proximal development that include new agents that are not human and can create new content. Multimodality theory must also adapt to AI-generated modes that are opaque to the learner, as well as adding an understanding of the probabilistic, data-driven nature of GenAI outputs to critical multimodal literacy. Activity theory also helps to explain the systemic paradoxes that emerge: between the pace of production and the depth of learning, between algorithmic fluency and cultural authenticity, between individualised scaffolding by AI, and the sociality of language development. The productive way to deal with these contradictions is a large learning challenge for researchers and practitioners for the next ten years.

From an SLA perspective, GenAI can drive up and enhance the opportunities for input, output and interaction, but the quality of the opportunities is dependent on design. When used without criticality, it can become a shallow prompt and accept tasks with underactive noticing and restructuring. On the other hand, carefully developed protocols can involve students in evaluating, editing, combining, and explaining what AI is doing, while also building their metacognitive and critical digital literacy skills. The interaction hypothesis (Long, 1996) takes a new turn: negotiation of meaning can now take place with an interlocutor - an AI - whose answers are fluent but at times somewhat unpredictable or biased, thus inducing authentic communicative pressure.

5.1 The Integrated Multimodal Interaction Model of GenAI-Augmented Digital Storytelling (IMIM-GADS)

Synthesizing the theoretical backgrounds and empirical results, a new Integrated Multimodal Interaction Model for GenAI-Augmented Digital Storytelling (IMIM-GADS) is suggested. The model views language learning with DST as an ecological system consisting of four interdependent interaction loops embedded

within a sociocultural and technological interaction system.

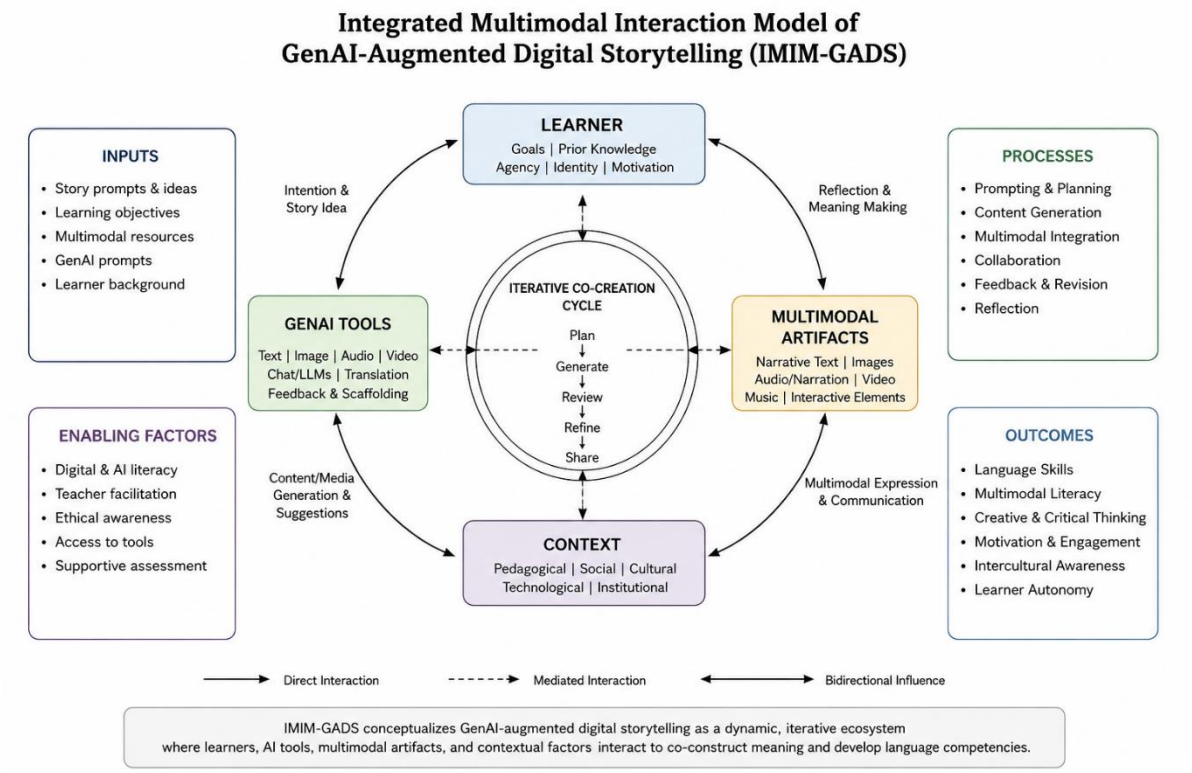


Figure 1: Integrated Multimodal Interaction Model for GenAI-Augmented Digital Storytelling (IMIM-GADS)

Institute for Excellence in Education & Research

5.1.1 Core Components

As shown in Figure 1, the Learner-Agent is at the center, with goals, prior knowledge, linguistic resources, identity investments, and prompt literacy impacting all interactions. The three major mediational resources surrounding the learner are the GenAI Agent (LLM and multimodal generative systems), the evolving storyboard, script, images, audio, video (Multimodal Artifacts), and the Sociocultural Context (peers, teacher, classroom norms, target culture, online audiences, institutional assessment regimes).

5.1.2 Four Interaction Loops

Loop 1: Narrative Scaffolding Interaction. The learner and GenAI carry out an iterative dialogic process to create plot, characters, setting, and thematic unity. The AI offers suggestions, elaboration, cultural and language equivalent options, and the learner evaluates, selects, and/or modifies and directs. This loop is mainly used to produce content, extend vocabulary, and structure discourses, and to develop critical thinking skills in relation to AI proposals.

Loop 2: Multimodal Orchestration Interaction. The learner uses the GenAI to create or edit specific modal elements (e.g., the image prompt, voice synthesis, music) and coordinates the use of linguistic, visual, auditory and interactive modes. The decisions made about what to show, rather than tell, how sounds evoke emotion, and how text and image interanimate are developed to achieve multimodal competence in designing. Decisions on what to show (non-show) and the way sounds create emotion, and the way text and image interact to create a meaningful message, develop multimodal competence in design and enhance semantic processing.

Loop 3: Feedback-Negotiation Interaction. Feedback comes from several sources: GenAI linguistic and content criticism, peer review, teacher feedback, and audience feedback post-

publication. The learner negotiates these inputs, accepting, rejecting and hybridizing what they wish to accept. This loop will play a key role in sensing, reframing and building self-regulation and critical AI literacy.

Loop 4: Identity Projection Interaction. The learner portrays elements of self, community, or an imagined self or community in the target language. GenAI can be used in this way to help express intricate personal or cultural stories that are outside the current range of productive skills, but that the learner must be able to express and keep them authentic. This loop involves investment, motivation and intercultural positioning.

5.1.3 Design Principles from the Model

As per IMIM-GADS, the five principles for practitioners are:

- The learner agency at the core of the learning experience remains non-negotiable; GenAI is a tool used to support, but not the end goal.
- Create cycles of generation-critique-revision-justification with specific guidelines to promote deep processing.
- Incorporate prompt literacy and critical analysis of AI-generated text as essential learning targets in addition to language learning.
- Balance the multimodal richness and cognitive load; give templates and progressive complexity.
- Incorporate ethical issues surrounding authorship, bias, privacy, and cultural representation into the curriculum on storytelling.

Table 4. Summary of IMIM-GADS Interaction Loops and Pedagogical Implications

Interaction Loop	Primary Learning Mechanisms	Key Pedagogical Moves
Narrative Scaffolding	Content generation; lexical expansion; discourse structure; critical evaluation of AI	Prompt workshops; AI suggestion critique protocols; learner-directed plot control
Multimodal Orchestration	Multimodal design competence; semantic deepening; aesthetic judgment	Mode-matching tasks; image–text interanimation analysis; progressive media complexity
Feedback-Negotiation	Noticing; restructuring; self-regulation; multi-source feedback integration	Revision justification logs; peer–AI–teacher feedback triangulation; delayed AI use
Identity Projection	Investment; motivation; intercultural positioning; authentic voice	Personal narrative prompts; cultural authenticity checks; public sharing with reflection

6. Implications for Practice, Policy and Research

6.1 Pedagogical Practice

Teachers should consider GenAI-augmented DST as an activity system that is designed to be implemented, and not a free-form technology experiment. These are some of the recommended practices: (a) using low-stakes prompt engineering tasks and AI critique before full story projects; (b) mandating process portfolios that include documentation of how revision tasks are divided between humans and AI; (c) incorporating explicit teaching of multimodal design principles and AI critical literacy, as well as designing process portfolios; (d) creating collaborative structures that maintain human-AI as well as human–human interaction; and (e) linking assessment with linguistic outcomes and higher-order human–AI interaction skills of orchestration, evaluation, and ethical judgment.

6.2 Assessment Innovation

The traditional product rubrics are inadequate. Process (revision history, justification for use of AI) and multimodal coherence, critical evaluation skills, and identity investment should be assessed. Promising directions are portfolio assessment, process-oriented rubrics, and learner self-assessment of the retention of agency. Balancing innovation with academic integrity is essential, and institutions must have clear policies regarding disclosure of AI use and acceptable use.

6.3 Research Agenda

Specific areas of focus are: longitudinal research on language development and identity trajectories; experimental analyses of various human-AI collaboration protocols; process-tracing and eye-tracking research of multimodal composition and AI-mediated composition; development and validation of instruments for multimodal communicative competence and AI-critical literacy; research on equity issues in under-resourced and multilingual settings; and ethical studies of data privacy, bias, and authorship in educational GenAI use. The cooperation between applied linguists, HCI researchers, AI ethicists and education academics will be crucial. Design-based research and research-practice partnerships are promising methodologies for refining the IMIM-GADS model and classroom implementations in authentic settings, in particular. Such relationships can help to ensure that theoretical ideas are translated into the real-world experiences of teachers and students, and that the type of thick, contextualized evidence captured by pure experimental designs is also generated.

7. Conclusion

The evidence-based findings show that the key pedagogical elements of DST, such as its use of engaging identity, its language modalities, its authentic purposes for language, and its process of creation and reflection, stand strong and, in many ways, are strengthened with the support of GenAI. The technology, however, presents new dilemmas to grapple with when it comes to agency, authorship, bias, assessment, and equity, issues that need to be carefully considered during pedagogical design, not determined by the technology itself. One answer to these tensions that IMIM-GADS offers is a framework that continues to place the learner in the centre of the learning process, but which clearly acknowledges the power of the new mediational means available. The Integrated Multimodal Interaction Model of GenAI-Augmented Digital Storytelling (IMIM-GADS) presents a theoretically and empirically informed framework for the design, research and evaluation of integrated multimodal practices. The model is centered on the agency of the learner, with explanations and practical suggestions through four dynamic interaction loops: narrative scaffolding, multimodal orchestration, feedback-negotiation, and identity projection. The language education community has a critical opportunity now to leverage these tools in ways that add to and not subtract from the human abilities that are fundamental to language learning: storytelling, meaning-making, and intercultural communication. This can only be done with methodological care, pedagogical purposefulness, ethical consciousness, and a commitment to sustained interdisciplinary communication; goals that the reviewers have tried to bring to bear in ways that are relevant to high-stakes scholarly engagement. Future research that tests, develops, and expands the model in various languages, learners, and environments will help to assess the huge potential of GenAI-augmented digital storytelling as a transformative tool in L2 writing instruction.

REFERENCES

Abdelrady, A. H., Akram, H., Mohamed, A. T. A., Abaalkhail, G. F., Osman, E. H. E., Mohamed, G. T. A. H., & Suliman, W. M. M. (2026). Harnessing AI for Pronunciation Development: Investigating the Impact of Elsa Speak on EFL Learners' Oral Proficiency. *World Journal of English Language, 16*(4), 311-321.

Abdelrady, A. H., Ibrahim, D. O. O., & Akram, H. (2025). Unveiling the Role of Copilot in Enhancing EFL Learners' Writing Skills: A Content Analysis. *World Journal of English Language, 15*(8), 174-185.

Akram, H., Abdelrady, A. H., Al-Adwan, A. S., & Ramzan, M. (2022). Teachers' perceptions of technology integration in teaching-learning practices: A systematic review. *Frontiers in psychology*, *13*, 920317.

Akram, H., Xi, W., & Li, S. (2026). Do emotional regulation behaviors matter in technology-integrated ESL pedagogical practices? A cross-sectional examination of teachers' attitudes. *Acta Psychologica*, 269, 107611.

Akram, H., Yingxiu, Y., Al-Adwan, A. S., & Alkhalsifah, A. (2021). Technology Integration in Higher Education During COVID-19: An Assessment of Online Teaching Competencies Through Technological Pedagogical Content Knowledge Model. *Frontiers in Psychology*, 12, 736522-736522.

Anderson, J., Chung, Y. C., & Macleroy, V. (2018). Creative and critical approaches to language learning and digital technology: Findings from a multilingual digital storytelling project. *Language and Education*, 32(3), 195–211.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.

Castañeda, M. E. (2013). "I am proud that I did it and it's a piece of me": Digital storytelling in the foreign language classroom. *CALICO Journal,* 30(1), 44–62.

Castañeda, M. E., Shen, X., & Claros Berlioz, E. M. (2018). This is my story: Latinx learners creating digital stories with VoiceThread. In J. P. Lantolf et al. (Eds.), *Handbook of sociocultural theory and second language learning*. Routledge.

Chapelle, C. A., & Sauro, S. (Eds.). (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell.

Chen, Z. (2025). *Generative AI technology and language learning. Interactive Learning Environments.* Advance online publication. Available at https://www.tandfonline.com/doi/full/10.1080/10494820.2025.2511248.

Chi, Y. (2025). Enhancing EFL/ESL learners' willingness to communicate through digital storytelling: A systematic literature review. *Educational Process: International Journal*, 14, Available at https://www.edupij.com/index/arsiv/79/787/enhancing-eflesl-learners-willingness-to-communicate-through-digital-storytelling-a-systematic-literature-review.

Cope, B., & Kalantzis, M. (2015). *A pedagogy of multiliteracies: Learning by design*. Palgrave Macmillan.

Creely, E. (2024). *Exploring the role of generative AI in enhancing language education*. Monash University Research Repository.

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics*, 35, 36–56.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156.

Godwin-Jones, R. (2023). Emerging technologies: AI and language learning. *Language Learning & Technology*, 27(2), 1–12.

Herrera Ramírez, Y. E. (2013). Writing skill enhancement when creating digital stories. *Colombian Applied Linguistics Journal,* 15(2), 166–183.

Hirsch, S., Macleroy, V., & Macleroy, V. (2020). *Multilingual digital storytelling: Engaging creatively and critically with literacy*. Routledge.

Hong, Q. N., et al. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. McGill University.

Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. *Written Communication, 22*(2), 224–261.

Hwang, W. Y., et al. (2016). Evaluating listening and speaking skills in a mobile game-based learning environment with dual-mode applications. *Computers & Education,* 88, 160–173. (Adapted context for DST oral outcomes).

Jalalzai, N. N., Akram, H., Khan, M., Kakar, A. K. (2025). Technology Readiness in Education: An Analysis of ICT Facilities in High Schools of Loralai, Balochistan. *Contemporary Journal of Social Science Review*, 3(3), 2835-2842.

Jewitt, C. (2013). *Multimodal approaches to learning. In C. Jewitt (Ed.), The Routledge handbook of multimodal analysis (2nd ed.).* Routledge.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Lambert, J. (2013). *Digital storytelling: Capturing lives, creating community (4th ed.).* Routledge.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Leong, A. C. H., Abidin, M. J. Z., & Saibon, J. (2019). Learners' perceptions of the impact of using digital storytelling on vocabulary learning. *Teaching English with Technology*, 19(4), 3–26.

Lim, N. Z. L., Zakaria, A., & Aryadoust, V. (2022). A systematic review of digital storytelling in language learning in adolescents and adults. *Education and Information Technologies*, 27, 6125–6155.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Ma, D., Akram, H. & Li, S. (2026). An integrated framework of ChatGPT adoption in higher education using TAM and UTAUT models. *Scientific Reports*. 16.

Ma, D., Akram, H., & Chen, I. H. (2024). Artificial Intelligence in Higher Education: A Cross-Cultural Examination of Students' Behavioral Intentions and Attitudes. *The International Review of Research in Open and Distributed Learning*, *25*(3), 134-157.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407–452). Oxford University Press.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–92.

Norton, B. (2013). *Identity and language learning: Extending the conversation (2nd ed.).* Multilingual Matters.

Ohler, J. (2013). *Digital storytelling in the classroom: New media pathways to literacy, learning, and creativity (2nd ed.).* Corwin.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, *37*, Available at https://www.bmj.com/content/372/bmj.n71.short.

Pardo, B. S. (2014). Digital storytelling: A case study of the creation, and narration of a story by EFL learners. *Digital Education Review*, 26, 74–84.

Popay, J., et al. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. ESRC Methods Programme.

Ramzan, M., Akram, H., & kynat Javaid, Z. (2025). Challenges and Psychological Influences in Teaching English as a Medium of Instruction in Pakistani Institutions. *Social Science Review Archives*, *3*(1), 370-379.

Ramzan, M., Awan, H. J., Ramzan, M., & Maharvi, H. (2020). Comparative Pragmatic Study of Print media discourse in Baluchistan newspapers headlines. Al-Burz, 12(1), 30-44.

Ramzan, M., Azmat, Z., Khan, M. A., & un Nisa, Z. (2023). Subject-Verb Agreement Errors in ESL Students' Academic Writing: A Surface Taxonomy Approach. In *Linguistic Forum-A Journal of Linguistics*, 5(2), 16-21.

Ramzan, M., Oteir, I., Khan, M. A., Al-Otaibi, A., & Malik, S. (2023). English learning motivation of ESL learners from ethnic, gender, and cultural perspectives in sustainable development goals. *International Journal of English Language and Literature Studies*, *12*(3), 195-212.

Ribeiro, S. (2015). Digital storytelling: An integrated approach to language learning for the 21st century student. *The EUROCALL Review*, 23(1), 34–41.

Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory Into Practice*, 47(3), 220–228.

Robin, B. R. (2016). The power of digital storytelling to support teaching and learning. *Digital Education Review*, 30, 17–29.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum.

Sweller, J. (2011). *Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), The psychology of learning and motivation (Vol. 55, pp. 37–76).* Academic Press.

Unsworth, L. (2008). Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In J. Coiro et al. (Eds.), *Handbook of research on new literacies*. Lawrence Erlbaum.

Vygotsky, L. S. (1978*). Mind in society: The development of higher psychological processes.* Harvard University Press.

Yang, Y.-T. C., & Wu, W.-C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & Education*, 59(2), 339–352.

Yang, Y.-T. C., Chen, Y.-C., & Hung, H.-T. (2022). Digital storytelling as an interdisciplinary project to improve students' English speaking and creative thinking. *Computer Assisted Language Learning*, 35(4), 840–864.

Yu, B. (2026). A systematic review of digital storytelling in English language learning and teaching. Educational Review. Advance online publication. Available at https://www.tandfonline.com/doi/full/10.1080/03055698.2026.2648654.

IEER
Institute of Excellence in Education & Research