

Experiential Narratives and Discourse Competence in EFL/L2 Education: A Systematic Review

Juan De Dios Marín Murillo^{1*}, Católica Luis Amigó Medellín²

^{1,2}Universidad Católica Luis Amigó Medellín, Colombia

¹Email: JuanDe12301@outlook.com

²Email: Medellin12351@outlook.com

Abstract

Discourse competence includes the ability to create and interpret communicative acts that are coherent, appropriate for a particular genre, and appropriately calibrated in terms of social context. To address this challenge, many teachers have turned to using experiential, narrative tasks such as digital storytelling. However, there has not been a thorough synthesis of what we know about how these types of tasks contribute to improving students' discourse competence. The results of a systematic review of the research on the topic of the application of experiential narratives to the formation of discourse competence in students. Among these 25 non-retracted articles were used in the synthesis analysis, and 2 retracted articles were stored as descriptive records but not as interpretative statements. The studies included in this review were selected based on searches of Scopus and Web of Science databases. The studies used in this review were all published between 2021 and 2026. The data collected from each of the studies was then examined for its methodological quality and relevance to the research question. Finally, the results of the studies were thematically synthesized into answers to four research questions. The four research questions addressed the theoretical frameworks that the author used to guide their research. This product is a preliminary result of the doctorate in education from the university.

Keywords: digital storytelling; discourse competence; EFL; narrative tasks; systematic review; intercultural communicative competence.

1. Introduction

1.1 Background and Rationale

Discourse competence is a major element of communicative language competence. It deals with the process of structuring language in coherent, cohesive, genre-acceptable and context-sensitive communication. These involve control of textual organization, pragmatic meaning, expectations of the audience, and appropriateness of text to sociocultural conditions. In EFL and second language classrooms, students might have an acceptable level of grammatical accuracy but still have difficulties in generating extended, well-organized discourse in speaking, writing or other multimodal communication. Discourse competence, hence, continues to be a significant teaching objective as well as a consistent pedagogical dilemma (Bardovi-Harlig, 1992; McCarthy, 1991).

Digital storytelling has become one of the key pedagogical activities in language learning in the last twenty years. It is more or less known as the composition of narrative texts with the help of digital media, typically incorporating written or spoken language, pictures, sound, video and other multimodal materials (Lambert, 2002; Robin, 2008; Quraishi et al., 2025). Digital storytelling in EFL and second language has been the focus of interest due to its opportunity to organize meaning in the narrative, use personal or lived experience, and express using a variety of semiotic modes. These aspects render it especially applicable to the process of discourse development since students must not only be able to use language correctly but also to structure ideas, to speak to audiences coordination of linguistic and visual meaning.

Experiential narratives have been explained in terms of a number of theoretical traditions in their pedagogical value. Socioculturally, Vygotsky (1978) focused on the meaning-making of the mediated

¹ Universidad católica Luis Amigó Medellín, Colombia

social activity, which defends the application of narrative as a means of learning and expression. In systemic functional terms, Halliday and Matthiessen (2004) demonstrated that communication achieves ideational, interpersonal and textual meanings concurrently, and thus, narrative production applies to discourse organization and social positioning. Recent intercultural methods, such as those that Byram (1997) discuss are, also propose that narrative activities can be used to help learners share identity, worldview, and culture. Collectively, these viewpoints suggest that experiential narratives can be used to add to discourse competence on a structural, pragmatic, multimodal, and intercultural level.

1.2 The Gap in the Literature

Despite the increased number of empirical studies on digital storytelling and the associated narrative tasks in the field of EFL and second language education, there are still two significant gaps in the field that are significant. First, discourse competence is commonly considered a general or implicit result and assessed by general scores of speaking or writing. With this trend, key differences between textual organization, pragmatic appropriateness, multimodal meaning making, and intercultural expression blur. Consequently, improvement is frequently reported in the literature, but the dimension of discourse competence that has improved is not clearly stated. Second, the review studies that have been done have investigated digital storytelling in language education more generally, yet have not narrowed down to discourse competence as the outcome of interest. Neither have they always used a clear quality appraisal procedure to assess the quality of the evidence available (Ateş, 2023; Ong & Aryadoust, 2023; Pérez Agustín & de la Peña, 2025).

1.3 Objectives and Research Questions

This systematic review fills these gaps with the synthesis of empirical evidence regarding experiential narrative tasks to be applied in order to develop discourse competence in EFL and second language situations. The review is dedicated to the studies that were published between 2021 and 2026:

RQ1: Which theoretical frameworks describe the relation between experiential narrative tasks and the development of discourse competence in EFL/L2 contexts?

RQ2: On which dimensions of discourse competence do empirical studies of DST-based language learning focus, and how are these dimensions operationalized?

RQ3: What does the current evidence show with respect to the effectiveness of experiential narrative tasks for the development of oral, written, multimodal and intercultural discourse competence across different educational levels and regional contexts?

RQ4: In which areas of methodological or thematic character does the existing literature currently have the biggest gaps, requiring priority in future research?

2. Methods

The systematic review was done based on the PRISMA 2020 statement and the PRISMA 2020 checklist of how to report systematic reviews (Page et al., 2021). The process of full reporting, such as the selection of the study, the criteria of eligibility, and the steps of its synthesis, was oriented on PRISMA 2020 and made the study transparent and reproducible. The review was not enrolled in PROSPERO due to the fact that the topic is not within the subject area of the platform of systematic reviews of health-related topics and research, but educational research as opposed to clinical outcomes. Before screening, a complete review protocol was created, which can be obtained by the authors on demand.

2.1 Inclusion/Exclusion Criteria

Table 1 displays the inclusion and exclusion criteria that were used throughout the selection process.

Table 1. Eligibility criteria.

Code	Inclusion criteria	Exclusion criteria
IC1/EC1	Peer-reviewed publications indexed in Scopus or Web of Science	Grey literature, theses, dissertations, book chapters, or non-peer-reviewed reports
IC2/EC2	Published between January 2021 and March 2026	Studies addressing narrative without an explicit focus on discourse or language competence

IC3/EC3	Addresses the use of experiential or personal narrative (DST, narrative retelling, oral storytelling, immersion-based narrative tasks, or telecollaborative narrative exchange) in educational contexts	Studies conducted exclusively in clinical or therapeutic contexts
IC4/EC4	Focuses on discourse competence as a primary outcome variable: oral discourse fluency, written discourse organization, narrative competence, multimodal discourse, ICC, argumentative writing, or productive skill development	Duplicate records across databases
IC5/EC5	Written in English or Spanish	Studies that address narrative as a research methodology rather than as a pedagogical intervention

Included in the review were only English and Spanish studies in Scopus and Web of Science because these databases are easily accessible, peer-reviewed, and consistent. Nevertheless, the use of these criteria can also create language, selection, and publication bias, as they will not include other sources of studies that are relevant.

2.2 Information Sources and Search Strategy

Two databases, Scopus and Web of Science, were searchable to identify literature through structured searches. The search strategy was based on Boolean combinations of three groups of concepts, namely, narrative form, educational context, and discourse competence. Search strings were transformed to the indexing capabilities of every database. A filter was used to reduce the searches to peer-reviewed publications as well as studies published in English or Spanish. The primary search was finished in February 2026, and an update search was conducted in March 2026 to include recent studies that were indexed.

Search String (Scopus / Web of Science):

```
("digital storytelling" OR "experiential narrative" OR "personal narrative" OR "narrative task" OR "storytelling")
AND ("EFL" OR "English as a foreign language" OR "L2" OR "second language" OR "language learning" OR "language education")
AND ("discourse competence" OR "communicative competence" OR "speaking skills" OR "writing skills" OR "oral discourse" OR "narrative competence" OR "intercultural competence" OR "productive skills")
```

2.3 Selection Process

The procedure of the study selection had the four steps of the PRISMA 2020 model. During the identification stage, 2,464 records were obtained in Scopus and Web of Science and added to Zotero to manage the records. Once duplicate removal had been done, 835 unique records were left to screen. During the screening phase, titles and abstracts were sifted through, and ASReview was employed to help in ranking the articles; the rest of the process was reviewer-led. This procedure eliminated 753 records, and only 82 articles were subject to complete text evaluation. During the eligibility step, the entire text of these 82 articles was assessed in terms of the set inclusion and exclusion criteria. Out of these, 55 articles were filtered out due to a lack of eligibility. During the inclusion phase, 27 articles were included in the review. Nevertheless, 2 studies that were retracted were not analyzed and only used to record the data descriptively, which left a total of 25 non-retracted studies to be used in the

final analysis. The entire selection pathway is also indicated in the PRISMA flow diagram in Supplementary Figure 1.

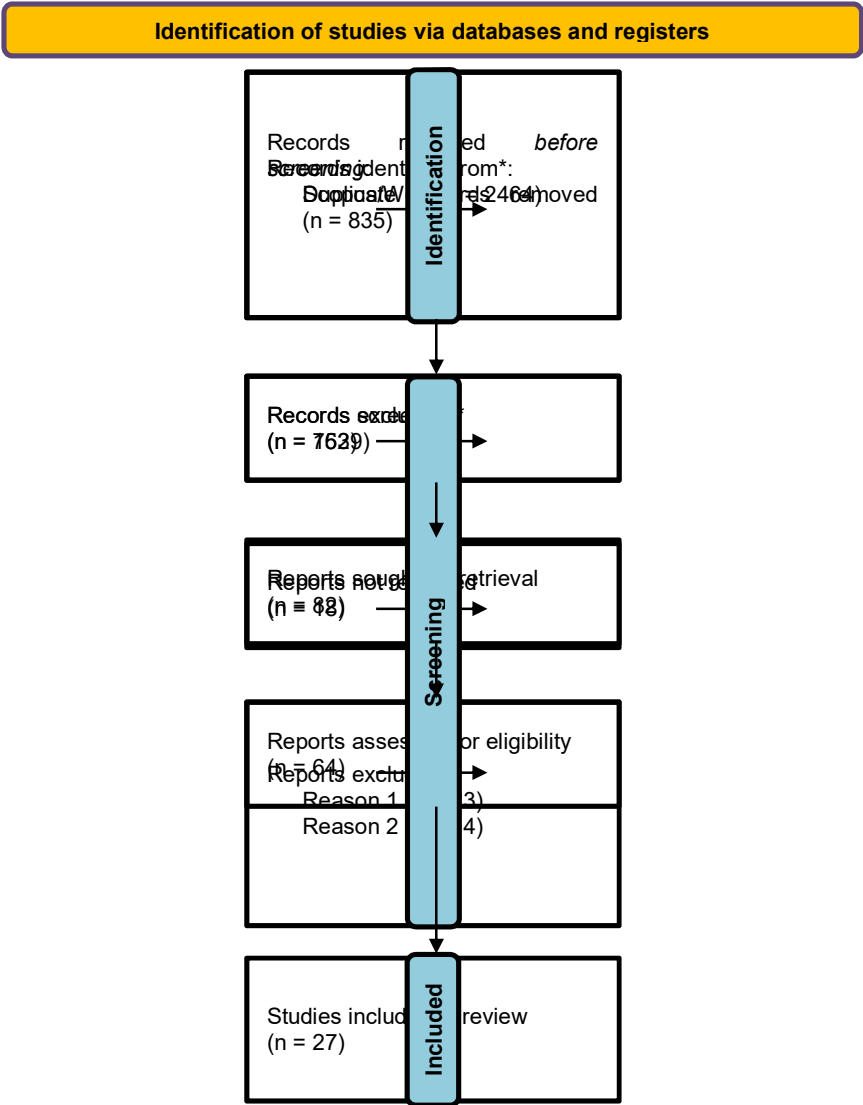


Figure 1. PRISMA 2020 flow diagram for the systematic review of experiential narratives and discourse competence in EFL/L2 education (Page et al., 2021).

2.4 Data Extraction

A pretested data extraction form was piloted on five studies and narrowed down before full use. The review retrieved bibliographic information, geographic and institutional setting, level of education, and profile of participants, methodological design, theoretical framework, type and operationalization of discourse competence, digital tools or platforms involved, significant findings, reported outcome measures or effect sizes, and quality appraisal findings of each of the included studies.

2.5 Quality Appraisal

The Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) was used to conduct the methodological quality appraisal. MMAT is not suggesting that one should compute a total numerical score. Rather, every study was evaluated regarding five design-specific criteria applicable to the methodological category of the study, such as the qualitative, quantitative, and mixed methods studies. To make the reporting transparent, the review summarized the number of criteria that each study fulfilled. The studies that met four or five criteria were said to have high quality, two or three

criteria were said to be of medium quality and zero or one criterion was said to be of low quality. These were merely descriptive summaries that were to be used to facilitate comparison among studies and not as official MMAT scores. Poorer-quality studies were included in the synthesis due to descriptive completeness, although this weight was less than that of other studies. Also, two withdrawn studies, Salem (2022) and Shelestova et al. (2023b), were only kept as records and not analyzed.

2.6 Synthesis Approach

The studies included were heterogeneous in terms of their methods and context; meta-analysis could not have been used. Rather, the review used a thematic synthesis methodology that uses a narrative synthesis as suggested by Thomas and Harden (2008). Findings of the study were coded line by line, and the descriptive themes and the higher-order analytical themes were created and mapped on the four review questions. To enhance the rigor of the methods, the coding decisions were verified between the extraction rounds and compared to the original results of the study. To enhance the development of categories and minimize subjective bias, a second reviewer reviewed a sample of coded studies and interpretations of them using thematic interpretation. The conflicting issues were solved by means of Discussion and numerous references to the original texts. Prejudice during coding was also reduced with the help of the pretested extraction form, audit trail of coding judgments, and cross-study comparison to make sure that the themes were not driven by preconceived notions but were based on the evidence.

3. Results

3.1 Study Selection and Characteristics

The systematic search indicated 2,464 records. Only 835 unique records were to be screened on titles and abstract after removing duplicates. A step of prioritization, facilitated by an AI and applied to aid a human-directed screening process, selected 82 articles to go through the full text. Among them, 27 articles were included in the descriptive review as they passed the eligibility criteria. Nevertheless, 2 studies were subsequently found to be retracted and were thus not included in the analysis, synthesis or all interpretations. In this light, the ultimate sample for analytical purposes was 25 studies that were not retracted. Two studies were later found retracted (Salem, 2022; Shelestova et al., 2023b) and were only retained alternative in descriptive reporting but not in any interpretation and synthesis of effectiveness.

The 27 studies were produced from 2021 to 2025, with 22 of these (or 81% of them) being developed in 2022 or later. This distribution shows that the existing evidence base regarding experiential narrative tasks in pandemic and post pandemic teaching environments is focused after 2022 (Alomari et al., 2023). Of the 14 countries of origin of the studies, the top four countries were Indonesia ($n = 5$), Türkiye ($n = 3$), China ($n = 3$), and Spain ($n = 3$). The largest proportion of studies (59%, $n = 16$) focused upon higher education. Secondary education was the focus for nine studies (33%), and there were two studies (7%) which focused upon primary education. Two of the studies included were systematic reviews, i.e., Ong and Aryadoust (2023) and Perez Agustin and de la Peña (2025). Any other form of review that was determined in the screening process was excluded at the screening stage since it did not pass the final inclusion criteria of scope, outcome focus and relevance of sample.

3.2 Quality Appraisal

The MMAT (Mixed Methods Appraisal Tool) quality appraisal of the 25 non retracted studies revealed the following distribution of study quality: high-quality studies that were assessed at four or five criteria ($n = 12$; 48%), moderate-quality studies that were assessed at two or three criteria ($n = 12$; 48%), and low-quality studies that were assessed at fewer than two criteria ($n = 1$; 4%). A full listing of all studies appraised with the MMAT can be found in Supplementary Table 2. Single-cohort pre-post design with no comparison group ($n = 6$) was the most frequently used design limitation, which makes it impossible to attribute the reported improvements causally.

Table 2. Characteristics of Included Studies ($n = 25$)

ID	Reference	Year	Context / Level	Design	Narrative Type	DC Target
S1	Ahdi et al. (2025)	2025	Higher Ed / Indonesia	Qualitative (narrative inquiry)	DST (student-created)	Creative thinking/discourse elaboration
S2	Amirinejad & Rahimi (2023)	2023	Primary / Iran	Quasi-experimental (pre-post)	DST (STEAM project)	English literacy/self-regulation
S3	Bai & Xian (2024)	2024	Higher Ed / China	RCT (pre-post + 1-month follow-up)	DST (personal narrative)	L2 speaking / self-regulation/anxiety
S4	Castillo-Cuesta et al. (2021)	2021	Higher Ed / Ecuador	Quasi-experimental (mixed)	DST (narrative writing)	EFL writing skills (grammar, vocabulary, organization)
S5	Dechsubha & Kanwivat (2023)	2023	Secondary / Thailand	Quasi-experimental (pre-post)	Storytelling (TCRS model: reading + role-play + storytelling)	Oral communicative competence/speaking fluency
S6	Gunawan et al. (2025)	2025	Secondary / Indonesia	Qualitative (multimodal analysis)	DST (factual autobiographical video)	Multimodal discourse / experiential meaning-making
S7	Kahanurak et al. (2023)	2023	Secondary & Higher Ed / Thailand	Theoretical review + reflection	DST (intercultural project-based)	Intercultural communicative competence (ICC)
S8	Liang & Hwang (2025)	2025	Secondary / Taiwan	Quasi-experimental	Human-robot collaboration DST (HRC-SDST)	Persuasive speaking / bodily-emotional awareness
S9	Liang et al. (2024)	2024	Higher Ed / China-Malaysia	Quasi-experimental (mixed)	Digital native storytelling (Chinese cultural narratives in English)	Native cultural awareness (NCA) / ICC
S10	Ölmez & Can Aran (2025)	2025	Secondary / Turkey	Quasi-experimental (pre-post)	DST (self-chosen topic)	EFL writing skills / FL anxiety
S11	Navas Romero (2025)	2025	Higher Ed / Spain	Survey (quantitative)	DST (student writing projects)	EFL writing skills/creativity/engagement

ID	Reference	Year	Context / Level	Design	Narrative Type	DC Target
S12	Ong & Aryadoust (2023)	2023	Multiple / Children	Systematic review (50 studies)	DST (various)	Written / oral / reading abilities
S13	Peña-Acuña & Navarro-Martínez (2024)	2024	Higher Ed / Spain	Mixed methods (quasi-experimental)	DST (multimodal stories in Spanish & English)	Originality / communicative competence/creativity
S14	Pérez Agustín & de-la-Peña (2025)	2025	Multiple / Spain	Systematic review	ST & DST (all types)	Oral and written linguistic competence (EFL)
S15	Qilin et al. (2025)	2025	Higher Ed / China-Colombia	Mixed methods (quasi-experimental + focus groups)	Telecollaborative narrative exchange	ICC (skills, awareness, knowledge, self-dimensions)
S16	Quay & Kano (2025)	2025	Secondary / Japan	Quasi-experimental (pre-post)	Immersion-based narrative (homestay storytelling)	Evaluative narrative competence/discourse complexity
S17	Roy & Bhavani (2025)	2025	Secondary / India	Sequential mixed methods (quasi-experimental)	DST (multimedia narrative)	Language development/student satisfaction
S18	Salem (2022) [RETRACTED]	2022	Higher Ed / Egypt	Quasi-experimental (3 groups)	DST (online & offline argumentative writing)	Argumentative writing / self-directed learning/learner autonomy
S19	Shelestova et al. (2023a)	2023	Secondary / Kazakhstan	Quasi-experimental (pre-post)	Narrative-based tasks (Flip/Flipgrid)	Productive skills (speaking + writing)
S20	Shelestova et al. (2023b) [RETRACTED]	2023	Secondary / Kazakhstan	Quasi-experimental (pre-post)	Storytelling (VoiceThread + Flipgrid)	Productive skills (speaking + writing) / confidence
S21	Silviyanti et al. (2022)	2022	Higher Ed / Indonesia	Qualitative (observation + interview)	Video storytelling (narrative)	Oral performance / communicative abilities

ID	Reference	Year	Context / Level	Design	Narrative Type	DC Target
					retelling: fables, legends, myths)	
S2 2	Supartini et al. (2024)	2024	Higher Ed / Indonesia (Tourism)	Quasi-experimental (R&D, exploratory mixed)	Storynomics (cultural narrative for tourism)	Speaking competence / professional communication
S2 3	Suryani et al. (2025)	2025	Higher Ed / Indonesia (Nursing)	Qualitative (narrative inquiry, semi-structured interviews)	DST (Storyboard That + PJBL)	Writing skills / narrative structuring / cognitive organization
S2 4	Syamsudin et al. (2025)	2025	Higher Ed / Indonesia	Qualitative (observation + interview)	Digital media storytelling (self-directed)	Speaking skills/anxiety/motivation
S2 5	Tatlı et al. (2022)	2022	Higher Ed / Turkey	Single case study (mixed: pre-post + interviews)	DST (student-created stories, 3 iterations)	Speaking skills/vocabulary / cognitive load

Table 2 indicates that there are some obvious trends within the analytical sample. To begin with, the evidence base is dominated by higher education as most of the research is carried out in universities, whereas primary education is poorly represented. This implies that the assertions on experiential narratives are based on the older groups of learners as opposed to early language learning situations. Second, the most frequent ones are quasi-experimental designs, and the number of randomized controlled trials is only one in the sample, which implies that the methodological strength of the field of study remains higher in exploratory intervention research than in high-rigor causal testing. Third, the most common type of narrative is digital storytelling and a smaller number of telecollaborative, immersion-based, and embodied narrative types are found that have more focused discourse outcomes, such as intercultural competence, persuasive communication, and evaluative narrative ability. Fourth, geographical location is clustered around Asia, particularly Indonesia, China, Thailand, and Turkey, which narrows down the generalizing contextually.

3.3 Theoretical Frameworks (RQ1)

Four theoretical orientations were clustered together. A constructivist orientation was also coded in 14 studies, when the authors directly described narrative learning as a socially mediated meaning-making, often referring to Vygotsky (1978) and Bruner (1991). Five studies employed Byram's (1997) Intercultural Competence (ICC) model (Kahanurak et al., 2023; Liang et al., 2024; Pérez Agustín & de-la-Peña, 2025; Qilin et al., 2025; Liang & Hwang, 2025). Two studies applied Systemic Functional Linguistics (Halliday & Matthiessen, 2004) as their primary analytic frame (Gunawan et al., 2025; Liang & Hwang, 2025). Three studies relied on Self-Regulated Learning (SRL) theory (Zimmerman, 2002): Bai & Xian (2024); Amirinejad & Rahimi (2023); Shelestova et al. (2023a). Genre-based frameworks and pragmatic theories, both of which are also important to the study of discourse competence as a construct, were notably absent.

Table 3. Theoretical Frameworks in the Included Studies

Framework	Role in DST/Discourse Studies	Studies
Constructivism (Vygotsky, 1978; Bruner, 1991)	Knowledge is constructed through active experience and social interaction; narrative provides scaffolding for language.	S2, S4, S5, S10, S11, S17, S23, S25
Systemic Functional Linguistics (Halliday & Matthiessen, 2004)	Language achieves meaning through ideational, interpersonal, and textual metafunctions; multimodal texts extend these functions.	S6, S8
Byram's ICC Model (1997)	Intercultural competence comprises five savoirs: attitudes, knowledge, skills of interpreting, skills of discovery and critical awareness.	S7, S9, S15
Self-regulated Learning (Zimmerman, 2002; Efklides, 2011)	Learners plan, monitor, and reflect on language production; DST cycles of planning-creating-reviewing scaffold SRL.	S2, S3
Cognitive-Interactionist SLA (Long, 1996; Ellis, 2005)	Language acquisition is fostered through negotiated interaction, noticing, and input richness; immersion operationalizes these conditions.	S16
Embodied Social Cognitive Theory (Bandura; applied HRI)	Bodily sensorimotor experience, social observation, and enactment support language ego development.	S8
Argumentation Theory / Rhetoric	Discourse competence extends to persuasive organization, claim-evidence structure, and genre mastery.	S18 (excluded from effect claims)
Creativity Theory (Torrance, 1998; Guilford, 1967)	Four-component model: fluency, flexibility, originality, elaboration; DST operationalizes each component.	S1, S13
Connectivism (Siemens, 2005)	Learning occurs through networked connections; transmedia tools like Flip scaffold community-based discourse practice.	S19
Cognitive Load Theory (Sweller, 1988)	DST places intrinsic load on language production, but extraneous load from technical tasks; design must manage both.	S25
Storynomics / ESP Narrative Pedagogy (McKee & Gerace, 2018)	Culturally embedded narrative is used as the primary vehicle for professional-domain speaking competence.	S22

Table 3 demonstrates that constructivism is prevalent in the discipline, whereas discourse-specific frameworks are not utilized well. Despite the fact that constructivism describes narrative learning in general, it fails to explain clearly the cohesion, coherence, genre, and pragmatics. The little application of SFL, argumentation, and interaction-based models contributes to understanding the existence of research gaps that are never resolved.

3.4 Discourse Competence Dimensions and Operationalization (RQ2)

Four discourse competence dimensions were identified. Oral discourse fluency was targeted in 17 studies, typically operationalized through multi-component rubrics assessing fluency, vocabulary, grammar, pronunciation, and coherence. Written discourse organization was targeted in nine studies through rubrics addressing grammar, vocabulary, narrative structure, and mechanics. Multimodal discourse was the focus of three studies (Gunawan et al., 2025; Liang & Hwang, 2025; Peña-Acuña & Navarro-Martínez, 2024). Intercultural discourse competence was operationalized in five studies primarily through validated ICC scales. A major discovery is that of all the studies, five employed instruments with clear indicators of discourse level, including cohesion, coherence, discourse organization, genre control or pragmatic appropriateness. The majority of studies have considered discourse competence as a synthesis of linguistic accuracy and overall proficiency, which restricts construct validity.

3.5 Effectiveness of Experiential Narrative Tasks (RQ3)

3.5.1 Oral Discourse and Speaking Skills

In 16 out of 17 studies that aimed at oral outcomes, authors documented statistically significant or qualitatively noticeable improvements in speaking-related measures after exposing participants to experiential narrative interventions. Bai and Xian's (2024) randomized controlled trial involving a blinded evaluation of speaking performance provided the best evidence, demonstrating sustained L2 speaking gains across all IELTS dimensions at one-month follow-up. Liang & Hwang (2025) demonstrated that embodied human-robot collaboration DST produced significantly superior persuasive speaking performance compared to animation-based DST ($p < .05$). Additional gains were documented in secondary contexts (Dechsubha & Kanwivat, 2023; Shelestova et al., 2023a), higher education (Tatlı et al., 2022; Supartini et al., 2024), and across cultural communities (Syamsudin et al., 2025). Silviyanti et al. (2022) provided a critical counterpoint: anxiety reduction in recorded storytelling did not automatically transfer to live audience-facing performance.

3.5.2 Written Discourse and Writing Skills

Nine studies examined writing outcomes. Castillo-Cuesta et al. (2021), Ölmez & Can Aran (2025), and Suryani et al. (2025) reported significant improvements in writing, with gains particularly pronounced in narrative organization, grammar, and vocabulary. The meta-analytic evidence of Ateş (2023) confirms positive effect sizes for writing across studies. Navas Romero (2025) found that learners strongly preferred DST over traditional writing on measures of creativity and engagement.

3.5.3 Multimodal and Intercultural Discourse

Gunawan et al. (2025) demonstrated through SFL analysis that students' DST autobiography videos constructed meaning predominantly through expository, identifying-and-describing processes, with multimodal image selection enabling meaning-tracking beyond purely linguistic resources. Qilin et al. (2025) found that a 9-week China-Colombia telecollaboration programmed produced significantly greater ICC gains than conventional instruction ($p < .001$), with the strongest effects in the skills and awareness dimensions. Liang et al. (2024) found that digital native storytelling improved the attitudes dimension of native cultural awareness, specifically the appreciation and expression of one's own cultural heritage in the target language.

3.5.4 Narrative Competence and Implicit Learning

Quay & Kano (2025) found significant post-immersion gains in narrative complexity, evaluative device use, and expression diversity in Japanese junior high school learners. Regression analysis identified age of first English exposure as a strong predictor of evaluative narrative competence, suggesting that early language exposure shapes the capacity for sophisticated narrative discourse in ways that brief interventions cannot easily compensate.

3.6 Methodological and Thematic Gaps (RQ4)

Four convergent gaps emerge. First, a measurement gap: the persistent conflation of linguistic accuracy with discourse competence prevents the field from establishing whether DST tasks develop discourse-specific skills. Second, a longitudinal gap: no study beyond Bai & Xian (2024) examined sustained discourse gains. Third, a primary education gap: only two studies operated below the

secondary level. Fourth, a geographic gap: 52% of active studies originated from East and Southeast Asia. An emerging fifth gap concerns the embodied and kinesics dimensions of discourse competence, which are underexplored relative to their theoretical importance. The results of these findings demonstrate a promising field with the potential for high pedagogical outcomes and still existing theoretical and methodological flaws. The Discussion below gives an interpretation of these patterns by associating them with the choice of frameworks, the practice of measurement, and limitations in research designs.

4. Discussion

4.1 Theoretical Frameworks: A Constructivist Consensus with Uneven Depth (RQ1)

The corpus reveals a field united by a broad constructivist consensus, but one that is insufficiently differentiated at the level of the specific cognitive operations involved in discourse organization. Fourteen of 25 active studies invoke Vygotsky or Bruner as foundational theorists; however, few of them elaborate on how narrative tasks are scaffolding the specific discourse processes that are relevant to competence, e.g., sequencing, cohesion management, register choice and genre identification.

The five studies employing Byram's (1997) ICC model represent the most theoretically precise strand of the corpus, precisely because that model specifies distinguishable dimensions of intercultural discourse competence that can be independently targeted and measured. By contrast, the near-absence of genre-based frameworks (Martin & Rose, 2008; Rothery, 1996), textual coherence and cohesion theories (Halliday & Hasan, 1976), and pragmatic frameworks (Grice, 1975; Kasper & Blum-Kulka, 1993; Al-Ezzi & Al-Qudah, 2024) is theoretically inconsistent for a body of research claiming discourse competence as its dependent variable. The textual metafunction of Systemic Functional Linguistics, which describes the organization of texts as communicative wholes, is directly pertinent to discourse competence, but applied as an analysis frame in only two studies (Gunawan et al., 2025; Liang & Hwang, 2025).

4.2 The Measurement Problem: Conflating Accuracy with Discourse Competence (RQ2)

The most consequential finding of this review is the systematic conflation of linguistic accuracy with discourse competence in measurement instruments. Across the majority of the corpus, 'speaking skills' and 'writing skills' are operationalized through composite rubrics that aggregate grammatical accuracy, vocabulary range, pronunciation, spelling, and mechanics, without distinguishing sentence-level production from discourse-level organization (Al-Ezzi & Al-Qudah, 2024). Only five studies applied instruments with explicit discourse-level indicators, demonstrating that the measurement gap is not inherent to the domain but represents an under-investment in purpose-built assessment tools.

This conflation has material consequences for practitioners. When a study reports that 'DST improved speaking skills', it is unclear whether gains involved improved discourse coherence, or simply broader vocabulary and greater grammatical accuracy. The pedagogical interventions required to develop cohesion and narrative structure differ from those needed to improve grammar; without discriminating instruments, the field cannot provide actionable guidance (Darwish et al., 2025). The development and validation of discourse competence assessment instruments disaggregating cohesion, coherence, genre management, narrative structure, and pragmatic appropriateness is the field's most urgent methodological priority.

4.3 Effectiveness: Robust for Proficiency, Tentative for Discourse (RQ3)

The evidence for DST effectiveness on general language proficiency is consistently positive and methodologically robust at its upper end. The double-blind RCT of Bai & Xian (2024), the meta-analytic evidence of Ateş (2023), and the systematic review evidence of Pérez Agustín & de-la-Peña (2025) converge in supporting the claim that experiential narrative tasks improve oral and written language production across contexts. More cautious conclusions are warranted, however, when the claim shifts to discourse competence specifically, given the measurement limitations identified above. Two findings provide the strongest evidence for discourse-specific effects. Liang & Hwang (2025) demonstrated that embodied narrative tasks, performed through physical, robotically mediated interaction, produced superior persuasive discourse outcomes compared to video-based DST,

suggesting that the communicative context of narrative performance shapes discourse development in ways that recorded production cannot replicate. Qilin et al. (2025) demonstrated that telecollaborative narrative exchange, linking learners across language communities, produced significant, theoretically specified gains in intercultural discourse skills that conventional instruction did not. Together, these findings suggest that the most productive direction for the field is not better video-based DST, but more interactive, contextualized, and socially embedded narrative tasks.

The findings on affective mediation also warrant attention. Anxiety reduction is a consistent mechanism linking DST to improved oral production across the corpus. Bai & Xian (2024) quantified this effect in a high-stakes IELTS context; Syamsudin et al. (2025) documented it qualitatively. However, Silviyanti et al. (2022) demonstrate that anxiety reduction in the private recording context does not automatically transfer to live communicative performance, underscoring the importance of designing progressive narrative task sequences that build toward authentic, interactive discourse contexts.

The findings related to using native culture narratives for developing Intercultural Communication Competence (ICC), as noted by Liang et al. (2024) and Kahanurak et al. (2023), propose an effective turnaround to a commonplace belief in digital storytelling pedagogy. Instead of students' focus being solely on learning about the culture of the target language, when students are asked to tell stories of their own cultural experiences in the target language, they can activate the reflexive, self-identity affirming aspect of Inter-Cultural Discourse Competence. This can be explained by wider sociolinguistic explanations that explore language acquisition as identity work, where the narration of lived experience in the target language will reinforce investment and voice (Praveen et al., 2025). This association is to be viewed as an interpretive extrapolation of the corpus given and not as an empirical conclusion to the studies that have been reviewed.

4.4 Gaps, Priorities, and Future Research Directions (RQ4)

Future studies need to focus more on discourse-specific measurement development. Validated instruments should distinguish discourse constructs, including cohesion, coherence, genre control, pragmatic appropriateness, and narrative macro structure, and sentence-level accuracy. In the absence of this distinction, the effects of interventions cannot be discussed as the manifestations of discourse competence improvement but rather as an overall change in proficiency. Longitudinal evidence is also imperative since discourse competence is a process that grows over long durations and cannot be expected to stabilize under short-term interventions. The studies must track learners over a minimum of one school year, must also contain repeated performance measures and look at long-term sustenance on measures beyond immediate post testing. Simultaneously, the field needs to invest more in primary education situations and in geographically underserved areas due to the existing evidence being overrepresented in secondary and higher education and being overrepresented in East and Southeast Asia. Lastly, the comparative effectiveness designs must test narrative modalities, embodied interaction, role play, synchronous telecollaboration and hybrid digital physical forms, as compared to the video-based digital storytelling as the control.

4.5 Summary of Findings by Research Question

Table 4. Summary of findings and implications by research question.

RQ	Principal Finding	Implication
RQ1	Constructivism dominates; ICC and SFL are underused as discourse-specific frameworks.	Future designs should integrate genre-based and pragmatic frameworks alongside constructivist scaffolding.
RQ2	Discourse competence is predominantly operationalized as general language proficiency, obscuring discourse-level constructs (cohesion, genre, pragmatics)	Instruments explicitly targeting discourse-level features are urgently needed.

RQ	Principal Finding	Implication
RQ3	DST consistently improves oral and written production; embodied and telecollaborative modalities yield the strongest discourse-level effects	Privileged modalities should move beyond video-based DST toward embodied, interactive, and cross-cultural narrative tasks.
RQ4	Measurement conflation, absence of longitudinal data, primary education gap, geographic concentration, and underexplored embodied dimensions	Longitudinal designs, primary-level studies, geographically diverse samples, and discourse-specific measurement are the four research priorities.

5. Conclusion

This systematic review, conducted in accordance with PRISMA 2020 guidelines, synthesized evidence from 27 studies on the use of experiential narrative tasks for the development of discourse competence in EFL and L2 educational contexts. The empirical case for experiential narrative tasks is robust, but the theoretical specificity of that case requires strengthening. DST-based interventions consistently improve general language proficiency across contexts, levels, and methodological designs. The corpus does not yet establish with equivalent rigor that these tasks develop discourse competence specifically, the capacity to produce and comprehend coherent, genre-appropriate, pragmatically calibrated communicative acts. Addressing this gap requires both more precise theoretical frameworks (genre-based, pragmatic, systemic-functional) and purpose-built discourse competence instruments. The narrative task modality and social structure are significant determinants of discourse outcomes. The contrast between the consistently positive but moderate effects of video-based DST and the stronger, more discourse-specific effects of embodied and telecollaborative narrative tasks suggests that how learners engage in narrative, physically, socially and interactionally, matters as much as the narrative content itself. Pedagogical designs should move learners progressively from private, scaffolded production toward live, interactive, audience-facing communicative performance. The native cultural narrative is an underexplored but theoretically productive site for intercultural discourse development. The evidence converges in suggesting that narrating one's own cultural experience in the target language activates identity-affirming, reflexive dimensions of discourse competence that target-culture-focused approaches may not develop. This constitutes a reframing of experiential narrative not merely as a language practice vehicle, but as a site for the negotiation of communicative identity across cultural contexts.

5.1 Limitations

The above review has several limitations. The inclusion of only those studies that appear on Scopus and Web of Science potentially excludes studies from other regional databases. The lack of PROSPERO registration before the study was carried out further reduces the potential for transparency. The methodological diversity among studies also makes a meta-analysis impossible. As well as the fact that two studies that had been retracted before being reviewed in this report were part of the original corpus, the failure to include these studies in the initial literature search reflects the limitations of the literature search process to identify post-publication integrity issues. The five-year time frame of this review was effective but limited in scope and does not allow for the examination of longer-term development trends in the research area. The search might have been biased by the database and language used since only Scopus, Web of Science and English or Spanish publications were searched, thus not including the relevant studies that were indexed in other databases or published in other languages. This could have restricted the cultural and regional diversity of evidence base, restricted the generalizability of results, and influenced the validity of conclusions concerning world trends in experiential narrative research.

5.2 Practical Implications

To practitioners, the data indicate that experiential narrative activities can be best justified when incorporated as a continuous pattern of instruction as opposed to single activities. The task design is to take learners through gradual disengagement of private and scaffolded drafting to live and interactive, audience-facing performance, so that the lesser anxiety in the recording situations can lead to the transfer of this anxiousness to actual communication. Telecollaboration is capable of reinforcing intercultural learning where possible by developing actual communicative stakes and a mutual conversation. The evaluation must be consistent with the given result of discourse competence with the help of the tools that measure the discourse level aspects, i.e., narrative structure, use of cohesion, coherence, genre control, and use of pragmatic appropriateness directly, instead of focusing on grammar and vocabulary correctness.

5.3 Future Research Agenda

The five research priorities identified in Section 4.4, discourse-specific measurement, longitudinal designs, primary education, geographic diversification, and embodied modalities, constitute an integrated agenda for the next phase of this research programme. A sixth priority should be added: comparative effectiveness research systematically testing the relative impact of different narrative modalities on specific, theory-derived discourse competence dimensions. Without such comparative evidence, the field cannot determine which narrative formats, under which instructional conditions, for which learners, and at which proficiency levels, most effectively develop the specific discourse capabilities that constitute communicative language competence in its fullest theoretical sense.

References

- Ahmad, M., Haider, A. S., & Saed, H. . (2025). Assessing AI-driven dubbing websites: Reactions of Arabic native speakers to AI-dubbed English videos in Arabic. *Research Journal in Advanced Humanities*, 6(1). <https://doi.org/10.58256/64dz1c72>
- Ahdi, A., Drajiati, N. A., & Putra, K. A. (2025). Exploring the lived experiences of designing digital storytelling among EFL pre-service teachers: A focus on creative thinking. *Electronic Journal of Foreign Language Teaching*, 22(1), 117–133. <https://doi.org/10.56040/ahdp2217>
- Al-Ezzi, R., & Al-Qudah, I. (2024). English translation of verbal humour in Egyptian comedy films. *World Journal of English Language*, 14(1), 157-157. <https://doi.org/10.5430/wjel.v14n1p157>
- Amirinejad, M., & Rahimi, M. (2023). Integrating digital storytelling into STEAM teaching: Examining young language learners' development of self-regulation and English literacy. *International Journal of Technology in Education*, 6(4), 720–735. <https://doi.org/10.46328/ijte.551>
- Ateş, M. (2023). Dil eğitiminde dijital öyküleme yönteminin kullanılması: Sistemik derleme çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 59, 340–364. <https://doi.org/10.9779/pauefd.1101722>
- Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
- Alomari, M. A., Khabour, O. F., Alzoubi, K. H., & Aburub, A. (2023). The Impact of COVID-19 Confinement on Reading Behavior. *Clinical Practice and Epidemiology in Mental Health*, 19(1), Article e174501792304260. <https://doi.org/10.2174/issn-vxx-yymmdd-2022-42>
- Bai, Y., & Xian, H. (2024). Exploring the interplay of digital storytelling, L2 speaking skills, self-regulation, and anxiety in an IELTS preparation course. *Humanities and Social Sciences Communications*, 11, 1584. <https://doi.org/10.1057/s41599-024-04109-8>
- Bardovi-Harlig, K. (1992). Pragmatics as a part of teacher education. *TESOL Journal*, 1(3), 28–32.
- Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, 18(1), 1–21. <https://doi.org/10.1086/448619>
- Byram, M. (1997). Teaching and assessing intercultural communicative competence. *Multilingual Matters*.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/1.1.1>
- Castillo-Cuesta, L., Quinonez-Beltran, A., Cabrera-Solano, P., Ochoa-Cueva, C., & Gonzalez-Torres, P. (2021). Using digital storytelling as a strategy for enhancing EFL writing skills. *International*

- Journal of Emerging Technologies in Learning, 16(13), 140–154. <https://doi.org/10.3991/ijet.v16i13.22187>
- Darwish, N., Haider, A., Tannous, B., Rumman, R. N. A., Alantari, D., Saed, H., & Dagamseh, M. (2025). A reception study of AI-translated idioms and proverbs between Arabic and English. *Humanities*, 6(3). <https://doi.org/10.58256/k4d6pp20>
- Dechsubha, T., & Kanwivat, S. (2023). Effects of teaching communication through reading, roleplaying, and storytelling (TCRS) on English instruction of high secondary school learners. *Theory and Practice in Language Studies*, 13(2), 331–340. <https://doi.org/10.17507/tpls.1302.13>
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics*, Vol. 3: Speech acts (pp. 41–58). Academic Press.
- Gunawan, W., Humaira, S. A., Hardini, T. I., & Zifana, M. (2025). Emerging digital factual storytelling in English language learning: Investigating multimodal affordances. *Studies in English Language and Education*, 12(1), 312–327. <https://doi.org/10.24815/siele.v12i1.41596>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). Arnold.
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). Mixed methods appraisal tool (MMAT), version 2018. Registration of Copyright (#1148552). Canadian Intellectual Property Office.
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.
- Kahanurak, S., Dibyamandala, J., & Mangkhang, C. (2023). Digital storytelling and intercultural communicative competence through English as a foreign language for multilingual learners. *Journal of Curriculum and Teaching*, 12(1), 14–28. <https://doi.org/10.5430/jct.v12n1p14>
- Kasper, G., & Blum-Kulka, S. (Eds.). (1993). *Interlanguage pragmatics*. Oxford University Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Lambert, J. (2002). *Digital storytelling: Capturing lives, creating community*. Digital Diner Press.
- Liang, J.-C., & Hwang, G.-J. (2025). Embodied social agency in human-robot collaboration for promoting EFL learners' persuasive communication and bodily-emotional awareness: A functional linguistics perspective. *Education and Information Technologies*, 30, 24291–24335. <https://doi.org/10.1007/s10639-025-13747-7>
- Liang, X., Mat Said, S. H., & Wei, L. (2024). Influences of digital native storytelling on foreign language learners' native cultural awareness. *Forum for Linguistic Studies*, 6(6), 1073–1087. <https://doi.org/10.30564/fls.v6i6.7100>
- Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.
- McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge University Press.
- Murad, T., Assadi, J., & Badarni, H. (2023). Digital storytelling and EFL speaking skill improvement. *Journal of Language Teaching and Research*, 14(5), 1189–1198. <https://doi.org/10.17507/jltr.1405.06>
- Navas Romero, C. (2025). User-friendly digital tools: Boosting student engagement and creativity in higher education. *European Public & Social Innovation Review*, 10, 1–17. <https://doi.org/10.31637/epsir-2025-963>
- Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind*. Cambridge University Press.
- Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. Longman.

- Ölmez, E., & Can Aran, Ö. (2025). The effect of digital storytelling on high school students' English writing skills and foreign language anxiety levels. *Participatory Educational Research*, 12(3), 234–252. <https://doi.org/10.17275/per.25.42.12.3>
- Ong, C., & Aryadoust, V. (2023). A review of digital storytelling in language learning in children: Methods, design and reliability. *Research and Practice in Technology Enhanced Language Learning*, 18, 11. <https://doi.org/10.1186/s40692-023-00258-y>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Peña-Acuña, B., & Navarro-Martínez, Ó. (2024). The promotion of originality is perceived in two multimodal storytelling applications: Storybird and Scratch. *Education Sciences*, 14(1), 21. <https://doi.org/10.3390/educsci14010021>
- Pérez Agustín, M., & de-la-Peña Álvarez, C. (2025). A systematic review of the positive effects of storytelling on the linguistic competence of English as a foreign language. *Bordón. Revista de Pedagogía*, 77(1), 103–119. <https://doi.org/10.13042/Bordon.2025.106007>
- Praveen, N. B., Quraishi, M. A., & Mudliar, S. S. (2025). Securing Low-Power Edge AI: A Vulnerability Analysis and Cybersecurity Framework for Resource-Constrained Devices. *International Journal of Innovation Studies*, 9(2). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5588331
- Qilin, Y., Mahfoodh, O. H. A., & Pérez Rubiano, P. A. (2025). Building intercultural bridges: Chinese EFL learners' intercultural communicative competence through telecollaboration with Colombian peers. *SAGE Open*, 15(4). <https://doi.org/10.1177/21582440251403241>
- Quay, S., & Kano, M. (2025). Implicit foreign language learning: How early exposure and immersion affect narrative competence. *Education Sciences*, 15(10), 1382. <https://doi.org/10.3390/educsci15101382>
- Quraishi, M. A., Salam, A., Agarwal, A., & Naca, P. Q. (2025). Cyber-Physical Co-Design Reliability Framework for ASIL-D Automotive Sensor ECUs with Integrated Hardware–Software Fault Tolerance and Security. *Journal of Information Systems Engineering & Management*, 10(63s), 671–683. <https://jisem-journal.com/index.php/journal/article/view/13917>
- Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st-century classroom. *Theory Into Practice*, 47(3), 220–228. <https://doi.org/10.1080/00405840802153916>
- Rothery, J. (1996). Making changes: Developing an educational linguistics. In R. Hasan & G. Williams (Eds.), *Literacy in society* (pp. 86–123). Longman.
- Roy, A., & Bhavani, S. (2025). Enhancing English learning through digital storytelling in Indian schools. *International Journal of Web-Based Learning and Teaching Technologies*, 20(1). <https://doi.org/10.4018/IJWLTT.368149>
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- Shelestova, T., Nabiyeva, A., & Kalizhanova, A. (2023a). Flip for promoting English language learners' productive skills: Connectivist practices in Kazakhstan. *Journal of Educators Online*, 20(2). <https://doi.org/10.9743/JEO.2023.20.2.7>
- Silviyanti, T. M., Achmad, D., Shaheema, F., & Inayah, N. (2022). The magic of storytelling: Does storytelling through videos improve EFL students' oral performance? *Studies in English Language and Education*, 9(2), 521–538. <https://doi.org/10.24815/siele.v9i2.23259>
- Supartini, N. L., Sudipa, I. N., Pastika, I. W., & Simpen, I. W. (2024). Integrating storynomics into language education to enhance the speaking skills of tourism students in Bali, Indonesia. *Journal of Curriculum and Teaching*, 13(2), 169–180. <https://doi.org/10.5430/jct.v13n2p169>
- Suryani, N. Y., Nugroho, H., Syahbani, M. H., Husein, I. G., & Rifaat, A. A. (2025). Integration of technology in project-based learning on digital story-writing to help students' writing skills. *Asian Journal of Human Services*, 29, 318–337. <https://doi.org/10.14391/ajhs.29.318>

- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Syamsudin, Budianto, L., P., K. D., Susanto, D., R., A., K., A., & Firdousi, M. A. (2025). Digital media's role in overcoming anxiety, enhancing linguistic elements and fostering motivation for developing speaking skills. *International Journal of Evaluation and Research in Education*, 14(2), 1379–1388. <https://doi.org/10.11591/ijere.v14i2.31931>
- Tatlı, Z., Saylan, E., & Kokoç, M. (2022). Digital storytelling in an online EFL course: Influences on speaking, vocabulary, and cognitive load. *Participatory Educational Research*, 9(6), 89–112. <https://doi.org/10.17275/per.22.130.9.6>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45. <https://doi.org/10.1186/1471-2288-8-45>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2