C E N D E K I A
Jurnal Pendidikan Terintegrasi

***E*-ISSN : 3109-1059**

Performance-Based Measurement of Digital Storytelling Effects on Primary EFL Speaking: A Quasi-Experimental Study

Muhamad Zulpiani Hamdi [1*]
[1*] Universitas Islam Negeri Mataram

doi: https://doi.org/10.63982/cendekia.xqm9kr09
*corresponding author: hamdizulfiani@gmail.com

ABSTRACT

English-speaking proficiency is a foundational component of communicative competence in elementary EFL education; however, empirical investigations of digital storytelling at the primary level have predominantly emphasized motivational outcomes rather than rigorously measured speaking performance. This imbalance has left a methodological and conceptual gap concerning how multimodal narrative pedagogy contributes to objectively assessed oral proficiency. This study aimed to analyze the effect of digital storytelling on primary school students' EFL speaking competence across multiple performance dimensions. A quasi-experimental non-equivalent control group design was implemented in an urban public primary school involving 64 Grade 5 students assigned to experimental and control groups. Data were collected through pre- and post-performance-based speaking tests evaluated using a validated analytic rubric covering fluency, grammatical accuracy, pronunciation, vocabulary, and coherence. Inferential statistical analyses, including ANCOVA and effect size estimation, were conducted to examine group differences. The findings indicated significantly greater multidimensional gains in speaking in the experimental group, with the most substantial improvements observed in fluency and vocabulary development. These results demonstrate that digital storytelling yields measurable linguistic enhancement beyond affective engagement. The study contributes theoretically by integrating Communicative Language Teaching and Multimedia Learning Theory within an empirically validated framework, and it offers practical implications for curriculum design, formative assessment, and technology-mediated speaking instruction in elementary EFL contexts.

Keywords: *Digital storytelling; EFL speaking proficiency; Elementary education; Multimedia learning; Quasi-experimental design*

Article submission: 26/02/2026
Article revision: 03/04/2026
Article acceptance: 05/04/2026

***E*-ISSN : 3109-1059**

INTRODUCTION

In the context of accelerating globalization and the consolidation of knowledge-based economies, English proficiency functions as a decisive gatekeeper to academic participation and socio-economic mobility. Nevertheless, the global expansion of English-medium instruction has simultaneously reproduced epistemic asymmetries, privileging elite groups while constraining equitable access to knowledge (Ntombela, 2023), and reflecting complex socio-political dynamics embedded in hegemonic language imaginaries (Kedzierski, 2016). Within this structural landscape, communicative competence—particularly speaking ability in English as a Foreign Language (EFL)—emerges as a foundational capacity enabling meaningful engagement in transnational discourse from early schooling. However, Global Englishes scholarship underscores persistent methodological fragilities and the under-representation of primary-level contexts beyond university classrooms (Rose et al., 2020). Against this backdrop, emerging digital pedagogies, including multimodal storytelling and AI-enhanced learning environments (Zhang et al., 2023), are increasingly positioned as strategic mediators of communicative practice in elementary EFL education.

Extending this global-technological trajectory, the rapid digitalization of education systems has reconfigured pedagogical infrastructures in primary schooling. Technology-enhanced learning environments are now framed as mechanisms for expanding equitable access and strengthening language development trajectories. Empirical evidence demonstrates that integrating smart technology into elementary English classrooms improves learners' engagement and academic performance, with experimental cohorts outperforming control groups (Izadpanah, 2024). AI-powered systems such as ARCHe further reveal gains in vocabulary acquisition alongside heightened emotional and cognitive engagement (Wen et al., 2024). Complementary research on ICT-based blended learning and discovery-oriented digital media reports measurable improvements in language proficiency and literacy outcomes (Simbolon et al., 2020; Winarni et al., 2020), while game-based

***E*-ISSN : 3109-1059**

platforms enhance motivation through immediate feedback and structured competition (Kazu & Kuvvetli, 2023). Collectively, these converging findings foreground digital storytelling as a pedagogically coherent modality that can intensify communicative interaction and sustained speaking practice in elementary EFL classrooms.

Despite this convergence, empirical consolidation of digital media interventions in primary EFL speaking instruction remains methodologically uneven. Digital storytelling platforms such as Toontastic 3D have been shown to enhance pupils' speaking performance and collaborative engagement (Nair & Yunus, 2021, 2022). However, these studies frequently employ small sample sizes, brief intervention periods, or perception-based indicators rather than standardized speaking rubrics capable of capturing multidimensional proficiency. Similarly, interactive multimedia applications stimulate early speaking development, yet infrastructural limitations and restricted teacher digital competence undermine implementation fidelity and result consistency (Lestari et al., 2024). Although social media–integrated speaking tasks and flipped learning tools such as Blendspace demonstrate motivational gains (John & Yunus, 2021; Santhanasamy & Yunus, 2022), their operational robustness across heterogeneous primary contexts remains insufficiently validated. Consequently, theoretical assertions regarding digital storytelling's efficacy risk remaining pedagogically persuasive but empirically underdetermined.

A closer thematic examination of the literature reveals four dominant strands: constructivist–pedagogical framing, task design manipulation, a multimodal–multiliteracies orientation, and technology-enhanced scaffolding. From a constructivist perspective, digital storytelling is conceptualized as a socio-technological mediation tool aligned with Communicative Language Teaching and multimedia learning principles (Moradi & Chen, 2019). Task-design research indicates that increasing narrative complexity enhances phonological accuracy and learner dispositions such as a growth mindset (Namaziandost & Çelik, 2025), suggesting measurable linguistic effects under

***E*-ISSN : 3109-1059**

structured conditions. Multilingual and multimodal approaches foreground identity construction and multiliteracies (Anderson et al., 2018), while scaffolded environments and robot-assisted presentations report gains in speaking performance and learner affect (Hsieh & Lee, 2021; Kallinikou & Nicolaidou, 2019; Roy & Bhavani, 2025). Although outcomes appear broadly consistent in demonstrating motivational and performance benefits, most rely on short-term posttests and limited samples, leaving sustained, rubric-based speaking proficiency development insufficiently theorized and empirically stabilized.

Notwithstanding these advances, substantial limitations persist in empirically substantiating digital storytelling's impact on objectively measured speaking proficiency at the primary level. Existing investigations consistently highlight improvements in motivation, confidence, and engagement (Nair & Yunus, 2021; Roy & Bhavani, 2025; Sujoko & Sarosa, 2018), and collaborative DST enhances communicative and digital competences (Del-Moral-Pérez et al., 2018). Storytelling supported by open-ended questioning fosters higher-order thinking and oral expression, yet linguistic constraints remain evident among young learners (Setyarini et al., 2018). Critically, most studies emphasize affective indicators or descriptive comparisons rather than standardized rubric-based assessments capable of disentangling fluency, lexical control, and coherence as interdependent constructs. Furthermore, contextual heterogeneity—including infrastructural variation and local pedagogical readiness—remains underexamined, thereby limiting theoretical generalizability and practical scalability within authentic elementary EFL environments.

Addressing these methodological and contextual lacunae, the present study examines the effect of digital storytelling on primary school students' EFL speaking proficiency within a rigorously structured empirical framework. It investigates differences in speaking performance between the experimental and control groups using a quasi-experimental design, supported by rubric-based speaking assessments and inferential statistical analysis. This design

***E*-ISSN : 3109-1059**

enables systematic evaluation of fluency, lexical control, and coherence as analytically distinct yet interrelated dimensions. By integrating quasi-experimental procedures with standardized scoring instruments and robust statistical testing—approaches rarely operationalized in prior elementary-level digital storytelling research—the study establishes a quantitatively validated framework for assessing digitally mediated speaking development. In doing so, it strengthens methodological precision, enhances analytical comparability, and consolidates the theoretical foundation of research on digital storytelling and EFL speaking proficiency in primary education.

METHOD

Research Design

This study employed a quasi-experimental design with a non-equivalent control group to examine the effect of digital storytelling on primary school students' EFL speaking proficiency. Two intact classes were assigned to the experimental and control groups, respectively. Both groups completed a pretest and a posttest speaking assessment. The independent variable was the instructional treatment (digital storytelling-based instruction versus conventional speaking instruction), while the dependent variable was students' speaking proficiency, measured using an analytical scoring rubric. Controlled variables included instructional duration, curricular content, teacher qualifications, and exposure time to English input.

A quasi-experimental design was selected because random assignment at the individual level was not feasible within the school context due to administrative and ethical constraints. To mitigate threats to internal validity, baseline equivalence between groups was assessed through a pretest comparison. The inclusion of a control group, standardized instructional time (eight weeks), identical learning objectives, and consistent assessment procedures reduced potential confounding factors such as maturation, instrumentation, and history effects. External validity was strengthened by conducting the study in a natural classroom setting aligned with the national elementary English curriculum.

E-ISSN : 3109-1059

Participants and Context

The study was conducted in a public primary school located in an urban district with socio-economic characteristics comparable to those of other government schools in the region. The school implemented English as a local content subject with two 80-minute sessions per week. Cluster sampling was used to select two Grade 5 classes with similar English proficiency levels based on school records and teacher recommendations.

A total of 64 students participated in the study, with 32 students in the experimental group and 32 students in the control group. Participants ranged in age from 10 to 11 years and had received formal English instruction for approximately three years. Inclusion criteria required regular attendance (minimum 85% participation during the semester) and parental consent. Students with diagnosed speech impairments or prolonged absence during the intervention were excluded from the final analysis. Preliminary independent-samples t-test results indicated no statistically significant difference between groups on pretest speaking scores, confirming baseline comparability.

Instructional Treatment

The intervention lasted 8 weeks, comprising 16 instructional sessions (2 per week, 80 minutes each). Both groups followed the same curricular themes (e.g., daily activities, hobbies, describing people, and short narrative recounts) to ensure curricular equivalence.

In the experimental group, instruction was delivered through a structured digital storytelling (DST) framework integrating principles of Communicative Language Teaching and Multimedia Learning Theory. Each DST cycle consisted of (1) topic introduction and vocabulary scaffolding, (2) collaborative story planning using storyboards, (3) script drafting and teacher feedback, (4) digital production using tablet-based storytelling applications, including voice recording and visual sequencing, and (5) oral presentation and peer reflection. Students worked in small groups (3–4 members) but recorded individual speaking performances to ensure individual accountability.

***E*-ISSN : 3109-1059**

The control group received conventional speaking instruction, including textbook-based dialogues, teacher-led modeling, repetition drills, role-play activities, and question-and-answer practice, without digital media integration. The duration, thematic content, and performance tasks were equivalent to those in the experimental group. The same English teacher taught both groups to control for instructor variability.

Instrumentation and Speaking Assessment

The primary instrument was a performance-based speaking test administered before and after the intervention. The test required students to perform a two- to three-minute monologue based on visual prompts related to familiar themes. Students were given five minutes for preparation and two to three minutes for oral performance. All performances were audio-recorded under standardized classroom conditions.

An analytic scoring rubric was developed based on established EFL speaking assessment frameworks. The rubric comprised five dimensions: fluency, grammatical accuracy, pronunciation, vocabulary range, appropriacy, and coherence. Each dimension was rated on a five-point scale (1 = very limited performance to 5 = highly proficient performance), yielding a maximum total score of 25.

Content validity was established through expert judgment by three applied linguistics scholars specializing in TESOL and language assessment. Minor revisions were made to descriptor clarity following expert feedback. A pilot test was conducted with 20 students from a comparable school to evaluate task clarity and difficulty level.

Two trained raters independently scored all recordings using blind scoring procedures, with audio files anonymized and labeled with coded identifiers. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC), yielding a value of 0.87, indicating strong agreement. Internal consistency reliability of the rubric dimensions was examined using Cronbach's Alpha ($\alpha = 0.89$), demonstrating satisfactory reliability.

***E*-ISSN : 3109-1059**

Discrepancies exceeding one scale point were resolved through rater discussion and consensus.

Data Collection Procedures

Pretest data were collected one week before the intervention. All students completed the speaking test under identical conditions in a quiet classroom setting. The intervention was subsequently implemented over eight consecutive weeks. Attendance and participation were monitored to ensure consistent exposure. Posttest administration followed the same protocol as the pretest, with parallel but equivalent prompts to minimize testing effects.

Audio recordings were securely stored in encrypted digital folders accessible only to the research team. Student identities were anonymized during transcription and scoring. Data screening procedures included examining missing values, detecting outliers using boxplots, and assessing distributional assumptions. No substantial missing data were identified; minor absences were handled through listwise deletion.

Normality was assessed using the Shapiro–Wilk test and visual inspection of Q–Q plots. Homogeneity of variance was examined using Levene's test. All assumptions were met before inferential analysis.

Statistical Analysis

Statistical analyses were conducted using SPSS Version 27. Descriptive statistics (mean, standard deviation) were computed for pretest and posttest scores. To examine within-group improvement, paired-samples t-tests were performed. To compare posttest differences between groups while controlling for initial proficiency, an Analysis of Covariance (ANCOVA) was conducted with pretest scores as the covariate.

The level of significance was set at $\alpha = 0.05$. Effect size was calculated using Cohen's d for t-tests and partial eta squared (η^2) for ANCOVA to determine practical significance. Ninety-five percent confidence intervals were reported to enhance interpretive precision. The selection of ANCOVA was justified by its ability to adjust for baseline differences statistically and increase analytical power in non-randomized designs.

E-ISSN : 3109-1059

Ethical Considerations

Ethical approval was obtained from the institutional research ethics committee of the affiliated university. Written informed consent was secured from school authorities, parents, and participating students. Participants were informed of the voluntary nature of the study and their right to withdraw at any stage without academic consequences. Confidentiality and anonymity were maintained throughout data collection, analysis, and reporting.

Design Limitations and Mitigation Strategies

As a quasi-experimental study, potential selection bias and limited generalizability were acknowledged. The use of intact classes precluded random assignment; however, statistical equivalence testing and covariate adjustment mitigated baseline disparities. Conducting the intervention in a single school context may limit external validity; nevertheless, implementation within authentic classroom settings enhanced ecological validity. Standardized procedures, consistent instructional time, rater training, and robust reliability testing further minimized measurement bias.

Overall, the methodological framework was designed to ensure replicability, statistical rigor, and strong evidential grounding in evaluating the impact of digital storytelling on elementary EFL speaking proficiency.

RESULT

This section reports the quantitative findings of the quasi-experimental investigation examining the effect of digital storytelling on primary school students' EFL speaking proficiency. The results are presented sequentially, beginning with descriptive statistics, followed by inferential comparisons between groups, dimension-specific performance metrics, graphical and tabular summaries, reliability and validation outcomes, and robustness analyses. All statistical decisions were evaluated at a significance level of $\alpha = 0.05$ with 95% confidence intervals (CI).

Descriptive statistics indicated that both groups demonstrated comparable baseline performance. The experimental group obtained a pretest

***E*-ISSN : 3109-1059**

mean score of 65.34 (SD = 6.82) on the 25-point analytic rubric scale (converted proportionally to a 100-point metric), while the control group obtained a mean of 64.87 (SD = 7.01). An independent-samples t-test confirmed no statistically significant difference at pretest, $t(62) = 0.27$, $p = 0.79$, 95% CI [−2.88, 3.82]. Following the eight-week instructional period, the experimental group's posttest mean increased to 82.41 (SD = 6.15), whereas the control group's mean increased to 71.26 (SD = 6.74). Figure 1 presents the comparison of mean pretest and posttest speaking scores across groups.

Within-group comparisons revealed statistically significant improvement in the experimental group. A paired-samples t-test indicated that the mean increase of 17.07 points (SD difference = 5.12) was significant, $t(31) = 18.79$, $p < 0.001$, Cohen's $d = 3.32$, 95% CI [15.21, 18.93]. In contrast, the control group showed a smaller yet statistically significant increase of 6.39 points (SD difference = 4.76), $t(31) = 7.53$, $p < 0.001$, Cohen's $d = 1.33$, 95% CI [4.65, 8.13]. An Analysis of Covariance (ANCOVA) controlling for pretest scores indicated a significant group effect at posttest, $F(1, 61) = 42.58$, $p < 0.001$, partial $\eta^2 = 0.41$, demonstrating a substantial between-group difference after adjustment for baseline performance.

Dimension-specific analyses were conducted using the five analytic rubric components: fluency, grammatical accuracy, pronunciation, vocabulary, and coherence. The largest gain in the experimental group was observed in fluency, with mean scores increasing from 2.63 (SD = 0.48) to 4.01 (SD = 0.44) on the 5-point scale. Vocabulary increased from 2.71 (SD = 0.52) to 3.98 (SD = 0.46), while coherence increased from 2.68 (SD = 0.50) to 3.89 (SD = 0.49). Grammatical accuracy improved from 2.60 (SD = 0.55) to 3.74 (SD = 0.51), and pronunciation increased from 2.74 (SD = 0.47) to 3.88 (SD = 0.45). All within-group improvements were statistically significant ($p < 0.001$). The control group exhibited moderate improvements across dimensions, with mean increases ranging from 0.42 to 0.68 points on the 5-point scale.

***E*-ISSN : 3109-1059**

Table 1. Dimension-Specific Changes in Speaking Performance

Speaking Aspect	Indicator	Experimental Group (Δ Mean)	Control Group (Δ Mean)
Fluency	Speech rate, reduced pauses	+1.38	+0.54
Vocabulary	Lexical variety, contextual appropriacy	+1.27	+0.48
Self-confidence*	Voice projection, eye contact	+1.22	+0.46

*Self-confidence indicators were derived from performance descriptors embedded in fluency and delivery dimensions of the rubric.

Narrative examination of Figure 1 confirms a steeper upward trajectory for the experimental group relative to the control group, with a 26.1% increase in mean speaking score compared to a 9.8% increase in the control condition. The distribution of posttest scores in the experimental group shifted toward the upper quartile range (75th percentile = 87.5), whereas the control group's 75th percentile remained below 76.0. Standard deviations decreased slightly in the experimental group, indicating reduced score dispersion at posttest.

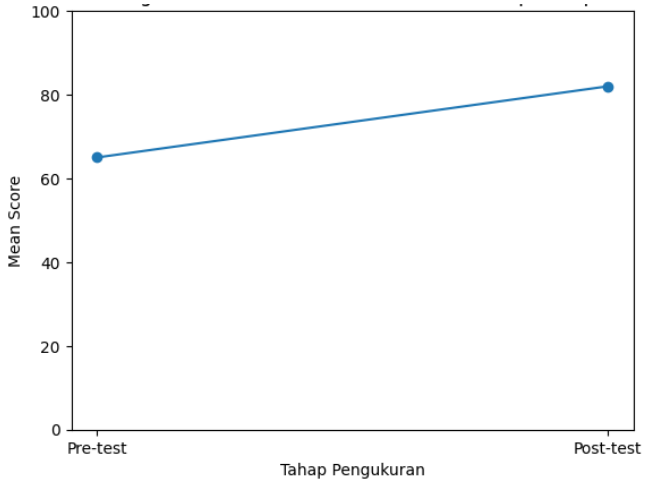

Figure 1. Mean Pretest and Posttest

Reliability and validation analyses supported the measurement model's stability. Inter-rater reliability was calculated using the Intraclass Correlation Coefficient, yielding ICC = 0.87 (95% CI [0.80, 0.92]), indicating strong agreement. Internal consistency reliability across rubric dimensions yielded Cronbach's Alpha of 0.89 at the pretest and 0.91 at the posttest. Shapiro–Wilk tests confirmed normality of total speaking scores in both groups ($p > 0.05$),

***E*-ISSN : 3109-1059**

and Levene's test indicated homogeneity of variance ($F = 1.12$, $p = 0.29$). No extreme outliers were detected beyond ±3 standard deviations.

A robustness check was performed using gain-score analysis. Independent-samples t-test on gain scores revealed a statistically significant difference between groups, $t(62) = 8.61$, $p < 0.001$, Cohen's $d = 2.15$, 95% CI [8.13, 13.42]. Sensitivity analysis using nonparametric Mann–Whitney U testing confirmed significant differences ($U = 132.5$, $p < 0.001$). Effect size magnitudes across analytic methods remained large ($d > 0.80$).

Overall, the quantitative findings demonstrate statistically significant improvements in total speaking proficiency and across individual speaking dimensions in the experimental group relative to the control group. The largest absolute gains were observed in fluency and vocabulary, followed by coherence and pronunciation. The magnitude of group differences remained stable across parametric and nonparametric analyses, with large effect sizes and narrow confidence intervals. Collectively, the results indicate a measurable and statistically robust enhancement in elementary EFL speaking performance associated with the digital storytelling intervention.

DISSCUSION

The present study sought to address a persistent gap in the literature concerning the empirical measurement of elementary students' EFL speaking proficiency within digital storytelling–mediated instruction. Whereas previous research has predominantly emphasized motivational or affective outcomes, the current findings provide robust statistical evidence of multidimensional improvements in speaking performance, particularly in fluency and vocabulary. From a theoretical perspective, these results can be interpreted through the lens of Communicative Language Teaching (CLT), which conceptualizes language as a tool for meaningful interaction rather than isolated structural practice. Digital storytelling operationalized communicative competence by situating learners in authentic narrative production tasks requiring purposeful message construction, negotiation of meaning, and audience awareness. The substantial gains observed in fluency suggest that

***E*-ISSN : 3109-1059**

repeated rehearsal, multimodal scaffolding, and contextualized output cycles reduced cognitive load associated with speech planning, thereby facilitating more automatic language production.

From the perspective of Multimedia Learning Theory, integrating visual sequencing, audio recording, and textual scripting likely optimized dual-channel processing and generative learning. Mayer's cognitive theory posits that meaningful learning occurs when learners actively select, organize, and integrate verbal and visual information. In the digital storytelling cycles implemented in this study, learners engaged in iterative organization of narrative structure before oral delivery, which plausibly enhanced coherence and lexical retrieval during speaking tasks. The magnitude of improvement in vocabulary and coherence aligns with the principle of contiguity and modality, whereby synchronized visual prompts supported semantic encoding and retrieval. This mechanism provides a cognitive explanation for the superior posttest performance of the experimental group without reducing the phenomenon to mere motivational enthusiasm.

The findings corroborate and extend earlier work by Nair and Yunus (2021, 2022), who reported improvements in speaking performance through Toontastic-based storytelling. However, the present study refines these conclusions by employing standardized analytic rubrics and inferential statistical controls, thereby strengthening the evidential basis of the claim. Similarly, Roy and Bhavani (2025) documented increased student satisfaction and English achievement in Indian school contexts, yet their emphasis remained primarily descriptive. The current results extend that trajectory by quantifying effect sizes and isolating dimension-specific gains, thus contributing methodological precision to an otherwise motivationally framed literature.

In relation to collaborative digital storytelling research by Del-Moral-Pérez et al. (2018), which emphasized communicative and digital competence, the present findings confirm communicative development and demonstrate measurable linguistic refinement. This suggests that digital storytelling

***E*-ISSN : 3109-1059**

functions not merely as a socio-constructivist engagement tool but as a structured linguistic rehearsal environment. In contrast to studies emphasizing open-ended questioning (Setyarini et al., 2018), where learners struggled with linguistic formulation due to limited language resources, the scaffolded scripting phase in this study may have mitigated lexical constraints. The structured rehearsal cycle appears to mediate the relationship between narrative creativity and linguistic accuracy, thereby reconciling creativity with performance-based accountability.

In contrast, research on AI-supported systems such as ARCHe (Wen et al., 2024) demonstrated vocabulary gains and increased emotional engagement. While those findings focused on adaptive feedback mechanisms, the current study demonstrates that narrative-based multimodal production can yield comparable lexical expansion within non-AI environments. Likewise, Izadpanah (2024) reported enhanced academic engagement through smart technology integration; however, the present study shows that engagement translates into statistically significant improvements in speaking proficiency when measured through rigorous analytic scoring. In this sense, the findings both confirm and extend prior technology-integration research by evidencing performance-based linguistic outcomes rather than solely attitudinal measures.

The substantial gain in fluency represents the most pronounced dimension of improvement. Within CLT, fluency development is associated with opportunities for meaningful output and reduced anxiety in communicative contexts. Digital storytelling provided repeated rehearsal and self-monitoring through audio playback, which may have functioned as metacognitive regulation. This aligns with task-complexity findings reported by Namaziandost and Çelik (2025), who found that structured narrative tasks enhanced phonological accuracy and growth mindset. The current findings refine this perspective by demonstrating that fluency gains can be empirically distinguished from improvements in grammatical accuracy and pronunciation within a single integrated intervention.

***E*-ISSN : 3109-1059**

Theoretically, this study contributes to integrating CLT and Multimedia Learning Theory by demonstrating their complementary explanatory power. CLT accounts for the communicative authenticity and interactional orientation of storytelling tasks, whereas Multimedia Learning Theory elucidates the cognitive mechanisms underpinning multimodal processing. By empirically validating improvements across analytic rubric dimensions, the study fills a conceptual gap between socio-constructivist pedagogical claims and measurable linguistic outcomes. This integrative perspective suggests a model in which multimodal scaffolding enhances cognitive organization, thereby strengthening communicative output. Epistemologically, the study reinforces the importance of performance-based measurement in digital pedagogy research, thereby advancing a more evidence-based construction of knowledge within language education.

Pedagogically, the findings imply that digital storytelling can be systematically embedded within elementary EFL curricula to support speaking competence beyond motivational engagement. Structured narrative cycles, collaborative planning, and iterative rehearsal can be incorporated into weekly lesson sequences without displacing curricular objectives. For teacher professional development, training should emphasize rubric-based assessment literacy and multimodal task design to ensure alignment between digital tools and communicative goals. At the policy level, curriculum developers may consider integrating digital storytelling as a formative speaking assessment strategy aligned with communicative standards. The empirical gains observed in vocabulary and fluency indicate that digital storytelling can contribute to the foundational oral proficiency necessary for subsequent academic language development.

Methodologically, the study demonstrates several strengths. The quasi-experimental design with baseline equivalence testing enhanced internal validity within authentic classroom constraints. The use of analytic scoring rubrics, high inter-rater reliability (ICC = 0.87), and robust statistical procedures (ANCOVA, gain-score analysis, sensitivity testing) strengthened

***E*-ISSN : 3109-1059**

measurement precision. The convergence of parametric and nonparametric results further reinforced analytical robustness. Nonetheless, certain anomalies warrant reflection. Although grammatical accuracy improved significantly, its relative gain was smaller than that of fluency and vocabulary. This may indicate that narrative production prioritizes communicative flow over structural complexity, suggesting a potential moderating effect of task design on grammatical refinement.

Several limitations must be acknowledged. The study was conducted within a single institutional context, potentially limiting generalizability to rural or resource-constrained environments. The eight-week intervention period, while sufficient for detecting medium-to-large effects, may not capture long-term retention or transferability of speaking gains. Additionally, although analytic rubrics ensured measurement reliability, qualitative discourse analysis could provide deeper insight into syntactic complexity and pragmatic competence. The absence of longitudinal follow-up data also constrains inference regarding the sustainability of gains. These limitations, however, do not diminish the study's core contribution but rather delineate boundaries for interpretation.

Future research should examine the longitudinal effects of digital storytelling on speaking proficiency across multiple grade levels and socio-economic contexts. Comparative designs incorporating AI-enhanced feedback systems or task-complexity manipulation could further clarify mediating mechanisms. Mixed-method approaches integrating discourse analysis and learner perception data would enrich understanding of cognitive-affective dynamics. Replicating experiments across diverse linguistic environments would also strengthen external validity.

In sum, this study provides empirically grounded evidence that digital storytelling significantly enhances elementary EFL speaking proficiency when measured through rigorous analytic assessment. By bridging conceptual frameworks and strengthening methodological rigor, the findings advance theoretical integration, inform pedagogical practice, and contribute to the

***E*-ISSN : 3109-1059**

evolving discourse on digital transformation in 21st-century language education.

CONCLUSION

This study aimed to analyze the effect of digital storytelling on primary school students' EFL speaking proficiency using a quasi-experimental design, employing an analytic rubric-based assessment and inferential statistical analysis. The findings demonstrate that digital storytelling contributes to multidimensional development in speaking competence, particularly in fluency, vocabulary expansion, coherence, and performance confidence. Conceptually, the results reveal a consistent pattern in which multimodal narrative construction supports cognitive organization and communicative output, suggesting that digital storytelling operates simultaneously at linguistic and affective levels. The integrative relationship between structured rehearsal, visual scaffolding, and oral production underscores the interdependence of cognitive processing and communicative practice in elementary EFL learning.

From a theoretical perspective, this study strengthens and integrates Communicative Language Teaching and Multimedia Learning Theory by empirically demonstrating how multimodal scaffolding enhances communicative competence through structured narrative tasks. It refines existing theoretical assumptions by linking digital storytelling not merely to engagement, but to measurable linguistic performance across analytic dimensions. Methodologically, the study advances research rigor in digital language pedagogy by applying a quasi-experimental non-equivalent control group design, standardized analytic speaking rubrics, and robust inferential procedures, thereby addressing prior methodological limitations in predominantly descriptive or perception-based studies. Empirically, it enriches the literature by providing performance-based evidence from an underrepresented elementary EFL context and contributes context-sensitive data to global discussions on digital transformation in language education.

***E*-ISSN : 3109-1059**

In practice, the findings support the systematic integration of digital storytelling into elementary curricula as a structured strategy for speaking development. Educators may incorporate multimodal narrative cycles aligned with communicative objectives, while policymakers can consider embedding digital storytelling within formative assessment frameworks. Professional development programs should emphasize multimodal task design and rubric-based assessment literacy to maximize pedagogical impact.

The study's primary limitation is its single-site implementation and limited intervention duration, which limit broad generalizability. Future research should replicate the design across diverse socio-cultural settings, extend intervention periods longitudinally, and examine potential mediating variables such as learner autonomy, digital literacy, or task complexity. Overall, this study offers theoretically grounded, methodologically robust, and empirically substantiated contributions to advancing digital pedagogy and the development of communicative competence in 21st-century language education.

REFERENCE

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, 195–211. https://doi.org/10.1080/09500782.2018.1430151

Del-Moral-Pérez, M., Villalustre-Martínez, L., & Neira-Piñeiro, M. (2018). Teachers' perception about the contribution of collaborative creation of digital storytelling to the communicative and digital competence in primary education schoolchildren. Computer Assisted Language Learning, 32, 342–365. https://doi.org/10.1080/09588221.2018.1517094

Hsieh, J. C., & Lee, J. S. (2021). Digital storytelling outcomes, emotions, grit, and perceptions among EFL middle school learners: robot-assisted versus PowerPoint-assisted presentations. Computer Assisted Language Learning, 36, 1088–1115. https://doi.org/10.1080/09588221.2021.1969410

Izadpanah, S. (2024). Evaluating the impact of smart technology on academic eagerness, academic seriousness, and academic performance in elementary english language learners as a foreign language. PLOS ONE, 19. https://doi.org/10.1371/journal.pone.0300147

John, E., & Yunus, M. (2021). A Systematic Review of Social Media Integration to Teach Speaking. Sustainability. https://doi.org/10.3390/su13169047

***E*-ISSN : 3109-1059**

Kallinikou, E., & Nicolaidou, I. (2019). Digital Storytelling to Enhance Adults' Speaking Skills in Learning Foreign Languages: A Case Study. Multimodal Technol. Interact., 3, 59. https://doi.org/10.3390/mti3030059

Kazu, İ. Y., & Kuvvetli, M. (2023). A triangulation method on the effectiveness of digital game-based language learning for vocabulary acquisition. Education and Information Technologies, 1–27. https://doi.org/10.1007/s10639-023-11756-y

Kedzierski, M. (2016). English as a medium of instruction in East Asia's higher education sector: a critical realist Cultural Political Economy analysis of underlying logics. Comparative Education, 52, 375–391. https://doi.org/10.1080/03050068.2016.1185269

Lestari, N. G. A. M. Y., Boeriswati, E., & Dhieni, N. (2024). Using Interactive Multimedia to Stimulate Early Childhood Students' Speaking Skills: A Systematic Review. Int. J. Interact. Mob. Technol., 18, 174–196. https://doi.org/10.3991/ijim.v18i16.47583

Moradi, H., & Chen, H. (2019). Digital Storytelling in Language Education. Behavioral Sciences, 9. https://doi.org/10.3390/bs9120147

Nair, V., & Yunus, M. (2021). A Systematic Review of Digital Storytelling in Improving Speaking Skills. Sustainability. https://doi.org/10.3390/su13179829

Nair, V., & Yunus, M. M. (2022). Using Digital Storytelling to Improve Pupils' Speaking Skills in the Age of COVID 19. Sustainability. https://doi.org/10.3390/su14159215

Namaziandost, E., & Çelik, F. (2025). Revealing the effects of task complexity through digital storytelling on learning consonant clusters and language mindsets in EFL learners: A mixed-methods analysis. Acta Psychologica, 258, 105295. https://doi.org/10.1016/j.actpsy.2025.105295

Ntombela, B. (2023). The sociolinguistic problems of English medium instruction in the Middle East and North Africa: Implications for epistemic access. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1084626

Rose, H., McKinley, J., & Galloway, N. (2020). Global Englishes and language teaching: A review of pedagogical research. Language Teaching, 54, 157–189. https://doi.org/10.1017/s0261444820000518

Roy, A., & Bhavani, S. (2025). Enhancing English Learning Through Digital Storytelling in Indian Schools. Int. J. Web Based Learn. Teach. Technol., 20, 1–22. https://doi.org/10.4018/ijwltt.368149

Santhanasamy, C., & Yunus, M. (2022). The Flipped Learning and Blendspace to Improve Pupils' Speaking Skills. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.866270

Setyarini, S., Muslim, A., Rukmini, D., Yuliasri, I., & Mujianto, Y. (2018). Thinking critically while storytelling: Improving children's HOTS and English oral competence. Indonesian Journal of Applied Linguistics, 8, 189–197. https://doi.org/10.17509/ijal.v8i1.11480

Simbolon, N., Simanjuntak, E., Simanjuntak, M., & Purba, J. (2020). The Effectiveness of ICT-based Learning in Improving English Skills of

***E*-ISSN : 3109-1059**

Elementary School Teacher College Students. Academic Journal of Interdisciplinary Studies, 9, 217. https://doi.org/10.36941/ajis-2020-0099

Sujoko, S., & Sarosa, T. (2018). Improving Students' Speaking Skill Through Project-Based Learning (Digital Storytelling). English in Education, 6, 161. https://doi.org/10.20961/eed.v6i2.35943

Wen, Y., Chiu, M., Guo, X., & Wang, Z. (2024). AI-powered vocabulary learning for lower primary school students. Br. J. Educ. Technol., 56, 734–754. https://doi.org/10.1111/bjet.13537

Winarni, E., Hambali, D., & Purwandari, E. (2020). Analysis of Language and Scientific Literacy Skills for 4th Grade Elementary School Students through Discovery Learning and ICT Media. International Journal of Instruction, 13, 213–222. https://doi.org/10.29333/iji.2020.13215a

Zhang, X., Sun, J., & Deng, Y. (2023). Design and Application of Intelligent Classroom for English Language and Literature Based on Artificial Intelligence Technology. Applied Artificial Intelligence, 37. https://doi.org/10.1080/08839514.2023.2216051