



Enhancing Translation and Subtitling Skills in EFL Contexts Through Digital Storytelling: A Mixed-Methods Study

Fitria Nugrahani

Politeknik Negeri Jakarta, Depok 16425, Indonesia

fitria.nugrahani@tik.pnj.ac.id

Abstract. This study introduces a pedagogical framework integrating digital storytelling (DST) with professional subtitling standards to enhance the translation and subtitling competencies of English as a Foreign Language (EFL) learners. By combining narrative creativity with technical precision, the approach bridges academic learning with real-world audiovisual translation practices. Using a mixed-methods design with 42 participants, quantitative assessments of subtitle timing accuracy and segmentation quality were conducted, referencing industry standards such as EBU-TT-D and ISO 20071-25. Results showed an 81% reduction in timing accuracy deviation and a 92% improvement in segmentation score (* $p < 0.001$). Qualitative analysis of reflective journals revealed that narrative ownership and iterative calibration during DST production fostered learner engagement and enabled millisecond-level synchronization. The findings indicate that DST not only facilitates language acquisition but also equips learners with practical, industry-aligned subtitling skills. This study contributes an evidence-based model for integrating DST into EFL curricula, offering a sustainable pathway to bridge translation training with professional media localization standards.

Keywords: Audiovisual Translation, Digital Storytelling (DST), EFL Learners, Language Learning Technology, Subtitling.

1 Introduction

Subtitling, particularly for animated content, is a demanding form of audiovisual translation (AVT) that requires both linguistic precision and technical accuracy. Professional standards emphasize frame-perfect synchronization with a tolerance of ≤ 100 milliseconds, optimal reading speeds between 12–17 characters per second, and segmentation that aligns with natural syntactic and prosodic boundaries [1], [2]. However, in English as a Foreign Language (EFL) education, subtitling has traditionally been taught as a form of textual translation, often reduced to simplified text-transfer activities that overlook these technical and perceptual requirements. For instance, common classroom practices involve static transcript matching or literal translation exercises without attention to timing, segmentation, or viewer processing

© The Author(s) 2026

E. Kurniawan et al. (eds.), *Proceedings of the 4th International Conference on English Studies in Indonesia (ICONESIA 2025)*, Advances in Social Science, Education and Humanities Research 1021,

https://doi.org/10.2991/978-2-38476-587-4_3

limits. This gap limits learners' ability to produce subtitles that are both professional and accessible, frequently resulting in segmentation errors, mistimed cues, and diminished viewer comprehension [3], [4], [5].

Previous EFL subtitling pedagogies have often fallen short for three reasons. First, they tend to prioritize linguistic accuracy over technical synchronization, leaving learners unprepared for the real-time demands of audiovisual media [3], [5]. Second, they rarely integrate authentic subtitling tools or industry standards (e.g., EBU-TT-D, ISO 20071-25), resulting in a disconnect between classroom practice and professional localization workflows [1], [2]. Third, such activities typically lack a creative or narrative component, which reduces learner engagement and motivation—key factors in sustained skill development [6], [7]. Without a pedagogical bridge that merges narrative engagement with technical training, EFL learners remain underprepared for professional subtitling environments.

Digital Storytelling (DST), which merges personal narrative with multimedia design, offers a promising solution to this challenge. Previous research has shown that DST supports vocabulary acquisition, enhances motivation, and fosters digital literacy through tasks such as video editing and voice-over recording [6], [7]. Beyond these linguistic benefits, DST also facilitates metacognitive development, as learners engage in cycles of self-review, timing adjustment, and iterative refinement [8], [9]. In EFL research, subtitling has been shown to improve multimodal literacy and pragmatic competence [10], [11], yet DST has not been systematically applied to the teaching of technical subtitling skills. This presents an opportunity to integrate narrative creativity with industry standards, preparing EFL learners for authentic translation tasks.

Building on these insights, the present study makes three contributions. First, it integrates DST with professional subtitling standards to train timing, segmentation, and reading speed within an EFL context. Second, it employs DST as a practical platform for calibrating subtitle duration and phrasing in real-time multimedia environments. Third, it investigates how learners develop segmentation strategies that align with speech rhythm and syntactic logic. Accordingly, the study addresses the following research questions:

1. To what extent does DST improve EFL learners' subtitling performance in timing and segmentation accuracy when benchmarked against professional standards?
2. What cognitive and linguistic strategies do learners employ during subtitle segmentation in DST tasks?
3. How can DST be implemented as a replicable pedagogical model to bridge the gap between academic translation practices and professional audiovisual subtitling?

By focusing on the intersection of creativity and technical precision, this study positions DST as a curriculum-ready solution for enhancing subtitling literacy and bridging the divide between language learning and media localization practice.

2 Method

This study employed a mixed-methods design to investigate the impact of Digital Storytelling (DST) on subtitling precision among EFL learners. Mixed-methods approaches are particularly suited for educational technology research because they combine the strengths of quantitative measurement with qualitative exploration, offering a holistic understanding of learner development [9]. The design integrated quantitative experimentation—focusing on frame-perfect timing, segmentation, and reading speed—with qualitative thematic analysis of reflective journals, thereby triangulating technical outcomes with learner perspectives.

The study involved 42 English as a Foreign Language (EFL) students enrolled in a multimedia study program at Politeknik Negeri Jakarta, where audiovisual translation is not included in the standard curriculum. Participants were selected through purposive sampling to ensure a baseline of limited subtitling experience, enabling clear measurement of learning gains. All participants underwent an introductory session on professional subtitling practices, followed by a four-week DST-based intervention.

The intervention was conducted across five phases: pre-test, DST-based training, post-test, reflective journaling, and integrative analysis. In the pre-test, participants subtitled a short animated clip to establish baseline performance metrics. During the four-week DST intervention, they created original digital stories and subtitled them using Aegisub, a professional subtitling tool. This stage emphasized compliance with broadcast norms such as reading speed, segmentation quality, and synchronization [3], [10]. In the post-test, learners subtitled a new animation clip of comparable difficulty, enabling measurement of progress through paired t-tests and ANOVA. Reflective journals were collected throughout the process, providing 210 journal entries that captured learners' strategies, challenges, and reflections.

Participants used Aegisub to create and refine subtitles, aligning their outputs with international subtitling standards. Timing accuracy was defined as subtitle cues occurring within ± 100 ms of speech onset or offset [8]. Reading speed was assessed based on the recommended range of 12–17 characters per second (CPS) [7], while segmentation quality was evaluated using ISO/IEC 20071-25 guidelines on syntactic and prosodic boundaries [4], [7]. These criteria ensured that the assessment reflected industry-level expectations rather than simplified classroom benchmarks.

The study evaluated four primary performance metrics pre- and post-intervention: timing accuracy (mean deviation from ideal in/out subtitle cues, measured in milliseconds), frame synchronization accuracy (percent of subtitles that precisely aligned with speech or scene changes), reading speed (average characters-per-second per subtitle, benchmarked against broadcast norms), and segmentation quality (percent of subtitles correctly segmented by phrase and syntax). Alongside quantitative measurement, participants maintained daily reflective journals during the DST process. These entries documented strategies such as the use of visual anchors, iterative revisions, and narrative-driven motivation. Braun and Clarke's six-phase framework for thematic analysis [9] guided the coding and interpretation process. Through joint displays, thematic patterns were integrated with performance metrics, enabling the identification of correlations between learner behaviors and technical improvements.

By combining experimental assessment with reflective inquiry, this method ensured both the reliability of performance measures and the depth of insight into learners' cognitive and affective processes.

3 Results and Discussion

The findings are organized into three primary areas: (A) quantitative improvements in subtitling performance metrics such as timing accuracy, frame synchronization, reading speed compliance, and segmentation quality; (B) qualitative insights derived from thematic analysis of student reflective journals, highlighting cognitive and emotional drivers of precision. These results are discussed in light of existing literature, with implications for subtitling pedagogy, curriculum design, and future research directions.

3.1 Quantitative Results: Subtitling Skill Improvements

Pre- and post-intervention comparisons showed significant gains across all subtitling metrics (Table I). Timing accuracy improved by 81%, with mean deviations reduced from 520 ms to 98 ms, placing most learners within the industry threshold of ± 100 ms [1] [12]. Frame synchronization errors decreased by 82%, while reading speed compliance rose from 65% to 92%, consistent with EBU recommendations of 12–17 CPS. Segmentation scores nearly doubled, reflecting stronger alignment with ISO/IEC 20071-25 standards [2]. These improvements indicate that DST-based subtitling tasks trained learners to meet professional-level requirements. Similar to findings by Talaván and Robin [8], [13], learners shifted from producing mechanically aligned subtitles to creating output that was linguistically coherent, temporally accurate, and accessible. The results extend earlier work by Sokoli [10], [11], who emphasized the pedagogical value of AVT, by demonstrating that systematic integration of DST accelerates learners' acquisition of technical subtitling literacy.

Table 1. Subtitling performance metrics (N = 42)

| Metric | Pre-Test (M ± SD) | Post-Test (M ± SD) | Improvement | p-Value |
|---------------------------|----------------------|-----------------------|------------------|---------|
| Timing Accuracy (ms) | 520 ± 148 | 98 ± 42 | 81% reduction | < 0.001 |
| Frame Sync Errors (count) | 8.2 ± 2.1 | 1.5 ± 0.8 | 82% reduction | < 0.001 |
| Reading Speed Compliance | 65% compliant | 92% compliant | +27% improvement | < 0.001 |
| Segmentation Score (1–5) | 2.4 ± 0.9 | 4.6 ± 0.7 | 92% improvement | < 0.001 |

Timing accuracy was measured by the mean deviation from ideal in/out points (measured in milliseconds). A threshold of ≤ 100 ms was considered "frame-perfect." The results show that 90% of participants achieved frame-perfect synchronization post-intervention, a significant jump from only 22% in the pre-test. Reading speed compliance was assessed based on the EBU recommendation of 12–17 characters per second (CPS). Post-intervention, 92% of participants produced subtitles within the

ideal range, compared to only 65% in the baseline. Segmentation scores were derived from a rubric adapted from ISO/IEC 20071-25 guidelines, evaluating whether line breaks matched natural linguistic units (clauses, phrases) and avoided visual clutter. Post-test performance showed a 92% improvement in average segmentation score ($t = 9.87, p < 0.001$).

3.2 Qualitative Insights: Cognitive Strategies and Learning Behaviors

Thematic analysis of 210 student journal entries revealed three recurring cognitive themes that contributed to measurable skill gains:

- **Visual-Auditory Synchronization:** 88% of participants reported using visual anchors (e.g., shot transitions, lip movements) to time subtitles precisely. One student noted: "I timed subtitle exits to scene cuts—visual anchors made sync intuitive." (Participant 19)
- **Iterative Calibration:** Learners engaged in repeated revision loops, with an average of 5.3 subtitle edits per clip. This helped them internalize timing intervals and identify subtleties: "I replayed my story 10+ times to fix 50ms gaps." (Participant 7)
- **Narrative Ownership:** The personalized nature of DST prompted emotional and creative investment, enhancing motivation to produce accurate, readable subtitles: "I cared about precision—it was my story peers would see." (Participant 33).

3.3 Integrated Analysis

Correlational evidence confirmed links between affective engagement and technical performance. Narrative ownership strongly correlated with timing accuracy ($r = 0.81$), while reliance on visual anchors predicted a 40% reduction in synchronization errors. Learners who demonstrated digital literacy with subtitling tools such as Aegisub showed faster calibration and higher compliance with CPS benchmarks [12]. These results illustrate how DST creates a convergence of creativity, technical training, and metacognitive growth.

3.4 Discussion

The study reveals that distinct elements of the DST process contributed differentially to subtitling skill development. Visual-auditory synchronization directly enhanced technical timing precision by providing concrete, scene-based cues for subtitle in/out points, moving learners beyond abstract timecode manipulation. Iterative calibration—facilitated by DST's cyclical creation process—was crucial for segmentation quality, as repeated revisions allowed learners to test and refine syntactic breaks for optimal readability. Finally, narrative ownership served as a motivational engine that increased persistence during technically demanding tasks, thereby improving overall compliance with reading speed and synchronization standards. This tripartite mechanism explains

why DST was more effective than traditional text-transfer exercises, which lack integrated visual, iterative, and affective dimensions.

Confirming, Extending, and Challenging Existing Knowledge

The findings of this study both confirm and extend prior research on AVT and technology-enhanced language learning. The significant improvements in timing accuracy and reading speed compliance align with earlier work by Sokoli [10] and Talaván [8], who highlighted the pedagogical value of active subtitling tasks. However, this study extends that research by demonstrating that such gains can be systematically achieved and measured against professional broadcast standards (e.g., EBU, ISO), a dimension often absent in classroom-based studies. More importantly, the results challenge the conventional separation of “creative” and “technical” skill development in translator training. Previous models have often treated narrative engagement and technical calibration as sequential or isolated processes [3][5]. In contrast, this study shows that in a DST framework, creativity and precision are synergistic: narrative ownership directly correlated with millisecond-level timing accuracy ($r = 0.81$), suggesting that emotional investment can drive technical excellence.

This finding contests the assumption that motivation and affect are peripheral to the acquisition of hard subtitling skills. Furthermore, the high segmentation scores achieved through iterative recalibration challenge the notion that syntactic awareness develops primarily through explicit grammar instruction. Instead, learners internalized segmentation principles through multimedia editing cycles—a process more aligned with situated and experiential learning theories [14]. This presents a meaningful departure from traditional, form-focused approaches to translation training in EFL contexts.

Bridging Creative and Technical Competencies

This study demonstrates how DST can transition subtitling from a passive translation task into an active, problem-solving process. By aligning subtitles with their own narratives, learners began to treat timing and segmentation not as constraints but as storytelling elements. This reflects Vygotsky’s theory of cognitive internalization, where external tools (e.g., timing software) become internalized strategies. Students intuitively adjusted their subtitles to match pacing, mood, and emotion [14].

Cognitive Load Optimization

The use of DST allowed learners to reduce extraneous cognitive load. Initial challenges with subtitle software (e.g., Aegisub) were mitigated as students progressed through cycles of creation and refinement. Rather than being overwhelmed by unfamiliar software and standards, learners approached timing as a natural extension of their story’s rhythm. One student noted: *"After my 3rd revision, I could feel 100ms gaps without checking the waveform."* (Participant 12)

Pedagogical Implications

This study offers a scalable model for integrating technical subtitling training into language learning. Enforcing industry standards (e.g., EBU/ISO) in the DST workflow prepared students for real-world translation tasks, led to high compliance (92%) in reading speed and segmentation [15]. Moreover, improvements in reading speed and segmentation have direct implications for accessibility, fluency, and viewer

comprehension as the approach bridges a crucial curriculum gap where audiovisual translation is rarely taught, especially in non-native English contexts.

4 Conclusion

This study demonstrated that Digital Storytelling (DST) can effectively enhance subtitling skills in EFL contexts by combining creative narrative production with professional technical standards. Learners achieved substantial gains across all measured metrics: an 81% improvement in timing accuracy (reducing mean deviation from 520 ms to 98 ms), an 82% reduction in frame synchronization errors, a 27% increase in reading speed compliance (from 65% to 92%), and a 92% improvement in segmentation quality. These quantitative gains were strongly supported by qualitative evidence, where narrative ownership showed a significant correlation with timing precision ($r = 0.81$), and iterative calibration led to an average of 5.3 revisions per subtitle clip. The study advances the understanding of DST in language education by moving beyond its established role in fostering motivation and digital literacy. It provides empirical evidence that DST serves as a viable pedagogical scaffold for technical, industry-aligned skill acquisition. By integrating professional standards (EBU-TT-D, ISO 20071-25) directly into the creative process, this research bridges a critical gap between narrative-based language tasks and the precision demands of professional audiovisual translation. It challenges the traditional compartmentalization of creativity and technical training, showing that affective engagement through personal storytelling can directly drive the mastery of millisecond-level synchronization and syntactic segmentation.

Several limitations should be considered. The study involved a single cohort of 42 multimedia students from one Indonesian polytechnic, which may affect the generalizability of findings to other EFL populations or educational contexts. Furthermore, the intervention relied on a specific tool (Aegisub), and its four-week duration limits insight into long-term skill retention.

Future research should address these limitations by expanding the sample size and diversity, incorporating learners from different linguistic backgrounds and academic disciplines. Longitudinal studies are needed to assess the durability of subtitling competencies acquired through DST. Additionally, research could explore the transferability of the DST model to other AVT modalities, such as dubbing or live captioning, and test its efficacy using different subtitling software platforms. Finally, cross-institutional implementations would help validate the scalability of this pedagogical framework and its adaptability to varied curricular constraints. By uniting creativity with technical rigor, DST represents a sustainable and replicable pathway for preparing EFL learners to meet the evolving demands of the media localization industry.

Acknowledgments. This study was conducted with the support of Politeknik Negeri Jakarta (PNJ). The authors gratefully acknowledge the funding provided by PNJ, which enabled the implementation of this research. We also extend our appreciation to the multimedia study

program, the 42 participating EFL students, and the faculty and technical staff whose cooperation and resources were essential to the success of this project.

References

- [1]. J. Díaz-Cintas and A. Remael, *Subtitling: Concepts and Practices*. Abingdon, UK: Routledge, 2021.
- [2]. ISO 20071-25:2017, "Information technology—User interface component accessibility—Part 25: Guidelines for captioning," International Organization for Standardization, Geneva, Switzerland, 2017.
- [3]. J. Lertola, "Audiovisual translation in EFL classrooms: A review of empirical studies," *J. Audiovis. Transl.*, vol. 3, no. 2, pp. 137–161, 2020.
- [4]. T. Alabsi, "Effects of adding subtitles to video via apps on developing EFL students' listening comprehension," *Theory Pract. Lang. Stud.*, vol. 10, no. 10, pp. 1191–1199, 2020.
- [5]. F. Chaume, *Synchronization in Dubbing: From Theory to Practice*. Abingdon, UK: Routledge, 2023.
- [6]. Y. Bai and H. Xian, "Exploring the interplay of digital storytelling, L2 speaking skills, self-regulation, and anxiety in an IELTS preparation course," *Humanit. Soc. Sci. Commun.*, vol. 11, Art. no. 1584, 2024, doi: 10.1057/s41599-024-04109-8.
- [7]. A. Mendoza-Escalante and J. C. [last name unknown], "Use of multimodal tools in teaching English as a foreign language," *Rev. Educ. Territ.*, vol. 6, pp. 71–86, 2016.
- [8]. N. Talavà, "Subtitling as a task and subtitles as support: Pedagogical applications," in *New Insights into Audiovisual Translation and Media Accessibility*, J. Díaz-Cintas, A. Matamala, and J. Neves, Eds. Amsterdam, Netherlands: Rodopi, 2010, pp. 285–299.
- [9]. J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Thousand Oaks, CA, USA: Sage, 2018.
- [10]. S. Sokoli, "ClipFlair: Foreign language learning through interactive revoicing and captioning of clips," in *Subtitles and Language Learning*, Y. Gambier, A. Caimi, and C. Mariotti, Eds. Bern, Switzerland: Peter Lang, 2015, pp. 127–148.
- [11]. B. Agulló Garcia, D. Orrego-Carmona, G. Spiteri Miggiani, R. Menezes, S. Valdez, S. Sokoli, and H. Pięta, "Indirect (pivot) audiovisual translation: A burning issue for research and training," *Perspectives*, vol. 32, no. 5, pp. 763–779, 2024, doi: 10.1080/0907676X.2024.2374643.
- [12]. A. Szarkowska and O. Gerber-Morón, "Viewers can keep up with fast subtitles: Evidence from eye movements," *PLOS ONE*, vol. 13, no. 6, Art. no. e0199331, 2018, doi: 10.1371/journal.pone.0199331.
- [13]. B. R. Robin, "Digital storytelling: A powerful technology tool for the 21st century classroom," *Theory Into Pract.*, vol. 47, no. 3, pp. 220–228, 2008.
- [14]. L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA, USA: Harvard University Press, 1978.
- [15]. B. R. Robin, "Digital storytelling in language education: A qualitative study on students' experiences," *J. Educ. Multimedia Hypermedia*, vol. 29, no. 4, pp. 321–340, 2020.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

