Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Designing Bilingual Storytelling Activities for Sardinian Language Engagement with AI

Gianni Fenu [1], Salvatore Carta [1,2], Alessandro Sebastian Podda [1], Mirko Marras [1], Marco Manolo Manca [1], Alessandro Giuliani [1,*] and Filippo Concas [1]

[1] Department of Mathematics and Computer Science, University of Cagliari; {fenu, salvatore, sebastianpodda, mirko.marras, marcom.manca, alessandro.giuliani, concas.filippo}@unica.it
[2] VisioScientiae S.r.l., Cagliari; salvatore.carta@visioscientiae.com
* Correspondence: Alessandro.giuliani@unica.it

Abstract: Minority and low-resource languages have limited digital resources. This limits their use in contemporary classrooms and can reinforce the view that they are less relevant or less usable. We propose an AI-supported framework that helps teachers build bilingual Italian-Sardinian activities around storytelling, listening comprehension, vocabulary, pronunciation practice, and interaction with story characters. The framework integrates speech recognition, machine translation, and speech synthesis in a teacher-mediated workflow, while envisaging an LLM-based conversational component for future interactive activities. The framework rests on a multidimensional view of engagement (cognitive, emotional, and social) and on technology as a cultural mediator. A typical activity starts with an Italian fable and introduces Sardinian step by step through subtitles, translation, and text-to-speech. It then moves to comprehension, guided production, pronunciation, and dialogue. We discuss how this approach may lower barriers for teachers who are not fluent in Sardinian, support attention to local linguistic variation, and link classroom practice to family and community knowledge. Future work will implement and evaluate the integrated framework in real classrooms settings, focusing on usability, perceived usefulness, learner engagement, pronunciation support, and cultural relevance.

Keywords: minority-language education; Sardinian; digital storytelling; learner engagement; artificial intelligence

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Digital technologies are nowadays widely used in education, making it easier to design interactive, multimodal, and participatory learning experiences (Huang et al., 2023). Speech interfaces, automatic translation, multimedia authoring tools, and conversational systems can support activities that combine listening, speaking, reading, writing, and visual content. These possibilities are not evenly distributed across languages. Most educational technology is built for high-resource languages, and minority and low-resource languages remain underrepresented in speech technologies, translation systems, and general-purpose AI tools (Joshi et al., 2020).

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

The consequences are not only technical. A language that is poorly represented in the digital environments that increasingly mediate learning and communication may come to be seen as less useful, less legitimate, or less compatible with current practices. Conversely, making a minority language visible and usable through digital tools can broaden the contexts in which learners meet and use it. Whether technology promotes linguistic inclusion or exclusion depends on which languages, varieties, and communicative practices it can support.

Sardinian is a relevant case of the aforementioned scenario. Although it is both a linguistic heritage and a living means of communication, its transmission to younger generations is uneven, and its presence in formal education is limited. Children may hear Sardinian from grandparents and relatives, or in local traditions, songs, and everyday expressions, but they have limited opportunities to use it systematically in other scenarios, such as at school. Indeed, teachers who want to bring Sardinian into the classroom also face practical hindrances: few age-appropriate materials, limited linguistic resources and technological tools, and varying levels of personal competence in Sardinian and its local varieties. Recent efforts have started to reduce this technological gap by developing dedicated speech resources for Sardinian and its linguistic varieties, providing a stronger basis for automatic speech recognition and related language technologies (Carta et al., 2025; Chizzoni & Vietti, 2024).

The challenge, then, is not just to supply translated content. Minority-language education requires that learners hear the language, understand it, pronounce it, use it in meaningful exchanges, and connect it to culturally relevant practices. Storytelling suits these requirements well, since it combines linguistic input with narrative structure, emotion, memory, imagination, and social interaction. Recent work on digital storytelling shows how story-centred and multimodal activities can support creative language use, oral production, identity expression, and learner participation (Kim et al., 2021; Macleroy et al., 2021; Nair & Yunus, 2021; Wong & Liem, 2022).

Based on these needs, this paper proposes an AI-supported framework for designing bilingual Italian-Sardinian activities. The main pedagogical scenario considered is storytelling through short stories and fables. The framework combines automatic speech recognition, machine translation, and speech synthesis, while also defining a planned LLM-based conversational component for future interactive activities. Together, these components are organized within a conceptual functional architecture intended to support listening comprehension, vocabulary development, pronunciation support, guided production, and contextualized interaction with story content. The framework is conceived as an educational mediator rather than as an autonomous tutor. The teacher selects and adapts the materials, validates linguistic outputs, determines when and how AI-generated resources are introduced, and situates them within a coherent pedagogical sequence. This teacher-mediated perspective is consistent with research emphasizing the role of instructional design and human–AI orchestration in language-learning environments (Wang et al., 2024; Weng & Chiu, 2023).

Our work is based on two main assumptions. First, learner engagement is a multidimensional process encompassing cognitive, emotional, and social dimensions rather than attention alone (Wong & Liem, 2022). Second, digital technologies are not neutral tools (Davoodi, 2024; Meighan, 2021; Pradhan & Dey, 2025): they mediate which languages, varieties, and forms of knowledge become visible,

accessible, and legitimate. Both assumptions are particularly relevant to minority-language education, where technology is closely connected with representation, identity, and linguistic continuity. Accordingly, the paper has three objectives:

1) to present the conceptual functional architecture of an AI-supported framework for bilingual Italian-Sardinian storytelling activities;
2) to illustrate an envisioned classroom workflow showing how different AI components can be combined under teacher mediation to generate reusable learning resources;
3) to discuss the educational and cultural implications, limitations, and future validation requirements of using AI-based language technologies to extend the presence of Sardinian in everyday classroom practice.

The contribution is therefore design-oriented and conceptual rather than evaluative: the framework defines the architecture, pedagogical rationale, and intended workflow, while its integrated implementation and educational effectiveness remain to be validated through future classroom-based studies.

2. Framework Design and Methodological Approach

This section presents the design rationale and methodological organization of the proposed AI-supported framework. It first summarizes the pedagogical and conceptual principles underlying the proposed approach, then describes the conceptual functional architecture and, finally, illustrates the classroom workflow through which these elements can be combined in bilingual Italian-Sardinian storytelling activities under teacher mediation.

2.1. Design Rationale

The proposed framework is grounded in a multidimensional view of learner engagement and in the assumption that digital technologies can act as educational and cultural mediators rather than merely as technical tools. Engagement is understood as involving cognitive, emotional, and social dimensions that emerge through interaction with tasks, teachers, peers, and learning environments (Wong & Liem, 2022; Meighan, 2024). In minority-language education, these dimensions matter especially because language learning is closely tied to identity, cultural memory, belonging, and the language's perceived legitimacy in contemporary contexts. In this context, the feasibility of resources depends on the maturity of the underlying language technologies. Automatic speech recognition, machine translation, and speech synthesis remain particularly challenging for low-resource languages because of the limited availability of annotated speech, parallel text, and high-quality speech corpora (Haddow et al., 2022; Pine et al., 2025; Slam et al., 2023), while using LLMs in language education introduces additional issues concerning output reliability, pedagogical integration, and human oversight (Law, 2024).

For Sardinian, the challenge therefore extends beyond vocabulary and grammar acquisition. Learners need opportunities to hear, understand, pronounce, and use the language in meaningful situations, while teachers need accessible resources that can support these activities even when ready-to-use materials are scarce or their own competence in a specific Sardinian variety is limited. Digital technologies may help reduce this gap, but their educational value depends on how they are embedded in

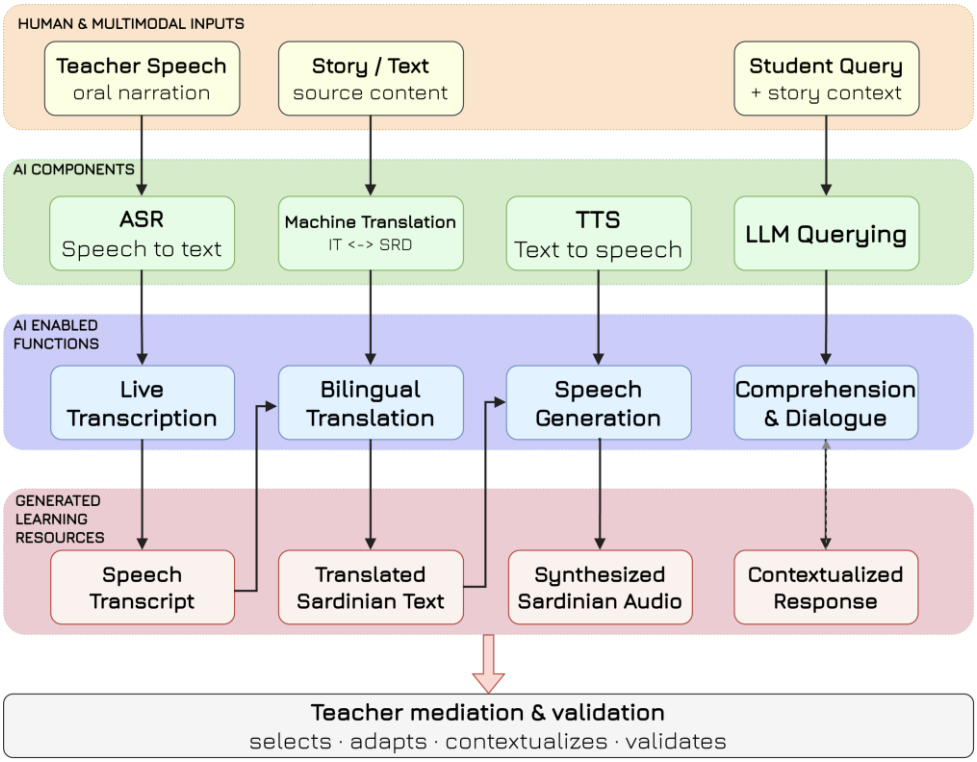
pedagogically meaningful activities rather than on the mere availability of AI functionalities.

For this reason, the framework adopts a teacher-mediated design. AI is not conceived as an autonomous tutor or an authoritative source of linguistic correctness. Instead, it provides computational functions that help the teacher prepare bilingual materials, support oral interaction, and adapt classroom activities. This perspective matters especially in a low-resource, linguistically variable context such as Sardinian, where automatically generated outputs may require linguistic and cultural validation.

2.2. Conceptual Functional Architecture

The proposed architecture is organized into four abstraction layers: **Human and Multimodal Inputs**, **AI Components**, **AI-enabled Functions**, and **Generated Learning Resources**. The architecture, depicted in Figure 1, should be interpreted as a conceptual design rather than as a description of a fully implemented system. Some components correspond to technologies already available or explored in previous work, whereas other functionalities, particularly the conversational LLM-based interaction, represent planned extensions to be implemented and validated in future work.

Figure 1. Conceptual functional architecture of the proposed framework.



The first layer describes the main inputs. **Teacher Speech** represents oral narration, such as the reading of a fable during a lesson. **Story/Text** represents

written source material selected or adapted by the teacher. **Student Query**, along with the relevant narrative context, represents the interactive input required by the framework's conversational branch.

The second layer comprises four AI components: **Automatic Speech Recognition (ASR)**, **Machine Translation**, **Text-to-Speech synthesis (TTS)**, and **LLM Querying**. These components provide distinct computational capabilities and are deliberately separated from their pedagogical use. ASR supports the conversion of oral narration into text. More specifically, the inclusion of ASR builds on recent advances in Sardinian speech technology and on the availability of dedicated speech resources covering multiple Sardinian varieties (Carta et al., 2025a). Machine translation supports Italian–Sardinian language mediation, TTS provides an oral realization of Sardinian text, and LLM querying provides the basis for contextualized interaction with story content. Previous work by our group has explored the use of machine translation and TTS for Sardinian language-learning activities, particularly through lexical translation and the oral realization of Sardinian content (Carta et al., 2025b).

At the functional level, these components enable **Live Transcription**, **Bilingual Translation**, **Speech Generation**, and **Comprehension and Dialogue**, respectively. The distinction between AI components and AI-enabled functions is central to the proposed design: the technologies themselves do not constitute learning activities, but become educationally meaningful when incorporated into teacher-designed classroom practices.

The corresponding outputs are represented as **Generated Learning Resources**: a speech transcript, translated Sardinian text, synthesized Sardinian audio, and a contextualized response. Outputs may also be reused across components. For example, an oral narration can be transcribed, the resulting text translated into Sardinian, and the translated version subsequently converted into speech. Similarly, a future LLM-generated response could be translated and synthesized when a spoken Sardinian answer is required.

Teacher mediation remains transversal to the entire architecture. The teacher selects and adapts source materials, evaluates whether generated outputs are appropriate for the targeted learners and linguistic context, and determines how each resource is incorporated into the activity. The architecture therefore separates processing-level automation from **human responsibility at the pedagogical level**.

2.3. Envisioned Classroom Workflow

Storytelling is the main pedagogical scenario the framework considers. A possible activity begins with a short story or fable that the teacher selects and adapts to the learners' age and educational objectives. The teacher may initially read the story in Italian, while ASR generates a textual transcription that can support comprehension and provide input for subsequent processing.

The teacher can then use the machine translation component to obtain a Sardinian version of the narrative and reproduce it through speech synthesis. Learners may listen to the Sardinian version while following the text or visual material associated with the story. This creates a gradual transition from familiar-language comprehension to exposure to Sardinian, rather than requiring learners to understand the entire target-language text immediately.

Subsequent activities may focus on comprehension and vocabulary. The teacher can select key words or expressions from the story, provide their Sardinian equivalents, and use synthesized speech to support repeated listening and pronunciation practice. The generated materials can progressively contribute to a shared class vocabulary resource or a small collection of reusable bilingual resources.

A further, currently envisioned extension concerns interaction with narrative characters. Learners could ask questions about the story or address a character directly, while an LLM processes the query alongside the story context and generates a contextualized response. The response could then be translated into Sardinian and reproduced through TTS. This functionality is currently conceptual and should therefore be considered a design objective rather than an evaluated capability of the present framework.

The educational purpose of this workflow is not to automate storytelling, but to increase the number of situations in which Sardinian can be heard, read, discussed, and reused in the classroom. The same architecture could subsequently be extended beyond fables to songs, poems, local stories, classroom routines, or materials related to other curricular subjects.

3. Discussion

The proposed architecture connects AI-based language technologies to pedagogical activities without making the technology itself the focus of the learning experience. Its main contribution lies in separating computational components, classroom-oriented functions, and educational resources. This distinction allows the reuse of individual AI services across different activities while preserving teacher control over their pedagogical role.

From an engagement perspective, the framework may support several complementary dimensions. The framework can encourage cognitive engagement through narrative comprehension, comparisons between Italian and Sardinian forms, vocabulary acquisition, and reformulation activities. Emotional engagement may arise from stories, cultural references, family expressions, and other content connected to learners' local environment. Social engagement may emerge through collective listening, repetition, discussion, and collaborative production of bilingual resources.

The framework may also help teachers who want to introduce Sardinian but do not feel confident enough to produce all linguistic materials independently. Translation and speech synthesis can potentially reduce the preparation burden and provide additional oral models. At the same time, automatically generated material should not be assumed to be linguistically neutral or universally appropriate. Sardinian shows substantial internal variation, and a model's output may differ from what a local community prefers. The teacher therefore remains responsible for validation and contextualization.

This consideration also motivates the framework's modular rather than fully autonomous design. A monolithic AI tutor would obscure the origin and transformation of generated content, whereas a modular architecture makes intermediate resources visible and, in principle, inspectable. A transcript can be checked before translation; translated text can be revised before synthesis; and an LLM response can be evaluated before being presented or transformed into speech.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

The conversational component raises additional considerations. Unlike transcription, translation, and speech synthesis, which can largely be interpreted as automated transformations, LLM querying introduces an open-ended interaction in which learner input dynamically shapes the resulting content. This makes teacher mediation particularly important, especially in activities involving young learners. Future implementations should therefore include mechanisms to constrain narrative context, control the type of generated response, and allow teacher validation before or during classroom interaction.

More broadly, the framework should be understood as a proposal for technology-supported language participation, rather than as a system for automated Sardinian teaching. Its potential value lies in expanding the contexts in which learners can encounter and use Sardinian within contemporary digital practices while maintaining connections with teachers, families, and community knowledge.

4. Limitations and Future Validation

The main limitation of the present work is that the proposed architecture has not yet been fully implemented or evaluated as an integrated educational system. In particular, the LLM-based conversational functionality remains a planned extension. Consequently, the paper does not claim demonstrated improvements in learning outcomes, engagement, pronunciation, or language use.

Future work should therefore proceed in two stages. The first concerns technical integration and validation of the proposed components, including the quality of Sardinian transcription, translation, speech synthesis, and contextualized LLM responses. Particular attention should be devoted to linguistic variation and the system's ability to support forms appropriate to the target classroom context.

The second stage should involve classroom-based evaluation with teachers and learners. Such evaluation should consider not only technical accuracy but also usability, perceived usefulness, teacher workload, learner engagement, comprehension, vocabulary acquisition, pronunciation support, and the perceived cultural appropriateness of generated resources. Qualitative feedback from teachers, families, and local language experts would be particularly valuable for assessing aspects that cannot be captured through technical metrics alone.

Future developments may also investigate mechanisms to represent linguistic variation explicitly, for example by allowing teachers to compare or select alternative Sardinian forms rather than presenting a single generated output as canonical.

5. Conclusions

This paper presented a conceptual functional architecture for AI-supported bilingual storytelling in Italian and Sardinian. The framework combines ASR, machine translation, speech synthesis, and a planned LLM-based conversational component within a teacher-mediated workflow that transforms human and multimodal inputs into reusable learning resources.

The proposal is intentionally modular and human-centered. AI components automate specific language-processing operations, while teachers remain responsible for selecting materials, validating outputs, contextualizing language use, and integrating generated resources into meaningful classroom activities. Storytelling offers an initial use case that lets learners move progressively from listening and

comprehension to vocabulary work, oral exposure, and potentially interactive dialogue.

At this stage, the framework should be regarded as a design proposal rather than a fully evaluated educational technology. Its contribution, therefore, lies in articulating how emerging language technologies may be organized and pedagogically mediated to support the presence and use of Sardinian in contemporary learning environments. Future implementation and classroom experimentation will be necessary to assess its technical feasibility, educational effectiveness, and cultural appropriateness.

References

Carta, S., Giuliani, A., Manca, M. M., Marras, M., & Piano, L. (2025a). *SardinianVoxes: A Speech Recognition Dataset for the Sardinian Languages*. Interspeech 2025, 1938–1942.

Carta, S. M., Fenu, G., Giuliani, A., Manca, M. M., Marras, M., Podda, A. S., & Pompianu, L. (2025b). *Enabling Sardinian Language Learning via Word Translation and Speech Synthesis based on Artificial Intelligence*. Journal of Inclusive Methodology and Technology in Learning and Teaching, 5(3).

Chizzoni, I., & Vietti, A. (2024). *Towards an ASR System for Documenting Endangered Languages: A Preliminary Study on Sardinian*. Proceedings of CLiC-it 2024, 214–220.

Davoodi, A. (2024). *Digital echoes of heritage: Toward a culturally balanced pedagogy in technology-enhanced bilingual education*. Journal for Multicultural Education, 18(1–2), 192–205.

Haddow, B., Bawden, R., Miceli Barone, A. V., Helcl, J., & Birch, A. (2022). *Survey of Low-Resource Machine Translation. Computational Linguistics*, 48(3), 673–732.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023*). Trends, Research Issues and Applications of Artificial Intelligence in Language Education*. Educational Technology & Society, 26(1), 112–131.

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). *The State and Fate of Linguistic Diversity and Inclusion in the NLP World*. ACL 2020, 6282–6293.

Kim, D., Yatsu, D. K., & Li, Y. (2021). *A multimodal model for analyzing middle school English language learners' digital stories*. International Journal of Educational Research Open, 2, 100067.

Law, L. H. L. (2024). *Application of generative artificial intelligence in language teaching and learning: A scoping literature review*. Computers and Education Open, 6, 100174.

Macleroy, V., Hackney, C., & Sahmland, S. (2021). *An Exploration of how Children's Language Learning can be Transformed when Teachers Place Creativity and Stories at the Centre of the Curriculum and Experiment with Digital Storytelling in the Classroom*. Modern Languages Open. https://doi.org/10.3828/mlo.v0i0.348.

Meighan, P. J. (2021*). Decolonizing the digital landscape: The role of technology in Indigenous language revitalization*. AlterNative, 17(3), 397–405.

Meighan, P. J. (2024). *Indigenous language revitalization using TEK-nology: How can traditional ecological knowledge and technology support intergenerational language transmission?* Journal of Multilingual and Multicultural Development, 45(8), 3059–3077.

Nair, V., & Yunus, M. M. (2021). *A Systematic Review of Digital Storytelling in Improving Speaking Skills*. Sustainability, 13(17), 9829.

Pine, A., Cooper, E., Guzmán, D., Joanis, E., Kazantseva, A., Krekoski, R., Kuhn, R., Larkin, S., Littell, P., Lothian, D., Martin, A., Richmond, K., Tessier, M., Valentini-Botinhao, C., Wells, D., & Yamagishi, J. (2025). *Speech generation for Indigenous language education*. Computer Speech & Language, 90, 101723. https://doi.org/10.1016/j.csl.2024.101723

Pradhan, U., & Dey, J. (2025). *Language, Artificial Education, and Future-Making in Indigenous Language Education*. Learning, Media and Technology, 50(2), 235–248.

Slam, W., Li, Y., & Urouvas, N. (2023). *Frontier Research on Low-Resource Speech Recognition Technology*. Sensors, 23(22), 9096. https://doi.org/10.3390/s23229096

Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). *Generative AI in education and research: A systematic mapping review*. Review of Education, 12, e3489.

Wang, F., Cheung, A. C. K., & Chai, C. S. (2024*). Language learning development in human-AI interaction: A thematic review of the research landscape*. System, 125, 103424.

Weng, X., & Chiu, T. K. F. (2023). *Instructional design and learning outcomes of intelligent computer-assisted language learning: Systematic review in the field*. Computers and Education: Artificial Intelligence, 4, 100117.

Wong, Z. Y., & Liem, G. A. D. (2022). *Student Engagement: Current State of the Construct, Conceptual Refinement, and Future Research Directions*. Educational Psychology Review, 34(1), 107–138