

TELÓ: AI-Driven Automatic Subtitling for the Promotion of the Performing Arts

Antoni Oliver, Silvia Rodríguez Vázquez, Manel Jiménez

Universitat Oberta de Catalunya (UOC)

{aoliverg, srodriguezvaz, manel.jimenez}@uoc.edu

Abstract

The TELÓ project provides an open-source framework for automated subtitling in the performing arts. Integrating state-of-the-art ASR and NMT, the system enables bidirectional translation between Catalan, Spanish, English and French. Designed for live performances, it provides synchronized captions for multiple devices, enhancing cultural internationalization and accessibility.

1 Project information

- Acronym: TELÓ
- Funding Agency: Generalitat de Catalunya
- Project Reference: PLG002/25/000052
- Duration: 6 months (02/02/2026 – 01/08/2026)
- Lead Institution: UOC

2 Introduction and Motivation

The performing arts in Catalonia represent a consolidated industry of high international value. However, the limited number of Catalan speakers among visiting populations and newcomers poses a significant barrier to the full dissemination and internationalization of theatrical works. While subtitling technologies have existed for years, they typically rely on manual synchronization during live performances and require labor-intensive prior translation.

TELÓ (Catalan for "curtain") leverages state-of-the-art Artificial Intelligence (AI) to automate this workflow, providing an open-source tool accessible to professional companies, small theaters, and amateur groups alike. The system supports bidirectional, automated subtitling between Catalan, Spanish, French and English, with an architecture designed for easy adaptation to additional language pairs.

© 2026 The authors. This article is licensed under a Creative Commons 4.0 licence, no derivative works, attribution, CC-BY-ND.

2.1 Context and Market Potential

The scale of the performing arts sector in Catalonia (787 spaces and 13,647 annual performances) highlights the language barrier: while non-verbal shows represent 8.33% of local performances, they account for 60.59% of international tours. Conversely, spoken Catalan performances drop from 60% locally to 12.38% abroad. TELÓ aims to reverse this trend by utilizing AI to make spoken-language drama viable for global audiences.

While Catalan is the primary focus, the system's multi-language capability is driven by immediate industry needs. At the time of writing, the first real-world implementation is scheduled for a French-language production to be subtitled simultaneously into Catalan and Spanish. Although the full calendar of performances for the pilot phase is still being finalized, this initial use case confirms the necessity of bidirectional support for the selected language pairs.

Beyond traditional theater, the modular architecture of TELÓ offers significant potential for adaptation to other live cultural contexts, such as audiovisual festivals, film screenings with live Q&A sessions, and international conferences. These environments share similar synchronization and translation challenges, making the TELÓ framework a versatile solution for the broader creative and cultural industries.

3 Project Objectives

The primary objective is to develop a AI-based system distributed under a free license to facilitate:

- Automatic Script Alignment: Synchronizing original scripts with existing translations.
- Domain-Adapted NMT: Generating high-quality translations tailored to the literary and dramatic genre.
- Real-time Subtitling: Delivering titles to projectors, mobile devices, and smart glasses.
- Accessibility: Providing same-language cap-

tions for the Deaf and Hard of Hearing (DHH) community.

4 Technical Framework

The project utilizes a robust stack of open-source technologies and research from the Aina Project and Grial Research Group:

- **Transcriber:** Employs speech-to-text models to process live audio. It can run local models, such as Whisper (Radford et al., 2023) and the Catalan-optimized Whisper (Hernández Mena et al., 2024), or commercial systems via API.
- **Synchronizer:** This module acts as the system’s core, aligning live transcripts with the pre-existing script and managing subtitle timing. It is capable of detecting actor improvisations; if a deviation from the script is identified, it automatically triggers the translator module to process the spontaneous dialogue in real time.
- **Translator:** Integrates the MTUOC framework (Oliver, 2025), supporting NMT and LLMs to translate both scripts and improvised dialogue. Future iterations of the project will focus on training translation systems specifically adapted to the linguistic registers and nuances of the performing arts.
- **Emissor:** Aggregates subtitles and generates web interfaces for broadcasting. Users can access independent language streams via any device (mobile phones, tablets, smart glasses). The system also supports direct projection onto physical screens.

This modular architecture facilitates adaptation to different user needs and viewing conditions. The system supports defining visual parameters like font size and color contrast depending on the device. To address theater policies, the interface includes a ‘Theater Mode’ with low-emission color schemes to minimize light pollution. Technical efficiency and low latency are ensured by the system’s ability to run on a local server via a dedicated Wi-Fi intranet, though it can also operate via standard web services. The use of a local dedicated network not only guarantees sub-second latency but also isolates the system from external interference, providing a robust environment for high-stakes live performances where timing is critical for the dramatic effect.

While the current prototypes use a cascaded ASR+NMT pipeline to leverage existing language-specific models, we intend to explore the integra-

tion of these components into a single end-to-end speech-to-text translation framework in future iterations.

5 Conclusions and future work

The TELÓ project provides an open-source, AI-driven solution for the performing arts. At the time of writing, the core architecture is implemented and preliminary tests have been conducted, with full-scale pilot trials scheduled for the final stage (concluding August 2026) to validate the system in professional settings.

Beyond its technical framework, TELÓ is a vital resource for cultural accessibility. Upcoming trials will prioritize feedback from actors, theater professionals, and the Deaf and Hard of Hearing (DHH) community to ensure the technology integrates seamlessly into the artistic flow. These insights are essential for refining the ‘Theater Mode’, ensuring the system acts as a non-intrusive bridge that enhances spectator immersion without distracting from the performance.

Future efforts will focus on enhancing robustness against acoustic unpredictability and improvisations. Designed for long-term evolution, the platform will be distributed under a GNU-GPL v.3 license on GitHub¹ to encourage community contributions. While currently focused on four languages, the methodology is designed for extrapolation to further language pairs, promoting social inclusion and the global positioning of local culture.

References

- Hernández Mena, Carlos Daniel, Carme Armen-tano Oller, Sarah Solito, and Baybars Külebi. 2024. 3catparla: A new open-source corpus of broadcast TV in Catalan for automatic speech recognition. In *Proc. IberSPEECH 2024*, pages 176–180.
- Oliver, Antoni. 2025. MTUOC server: integrating several NMT and LLMs into professional translation workflows. In *Proceedings of Machine Translation Summit XX: Volume 2*, pages 81–82.
- Radford, Alec, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *Proceedings of the 40th International Conference on Machine Learning*, pages 28492–28518.

¹<https://telouoc.github.io/>