



forthem.

Fostering Outreach
within European Regions,
Transnational Higher Education
and Mobility

Subtitling and Proofreading of FORTHEM Digital Academy Courses WP3 Report

1. Introduction

Within the framework of WP3 activities, FORTHM partners collaborated to improve the accessibility, multilingual reach and quality of educational resources available through the FORTHM Digital Academy (FDA). The activity focused on the **revision and proofreading of automatically generated subtitles and machine-translated subtitle versions** for a selection of Digital Academy courses.

The initiative was designed not only as a quality assurance activity for FDA learning materials but also as an educational opportunity for students. Through curricular and extra-curricular internships students from different FORTHM universities actively contributed to the improvement of subtitle quality while simultaneously developing practical skills in audiovisual translation, subtitling, multilingual revision, terminology management, post-editing and digital content accessibility.

The activity relied on a collaborative model involving academic supervisors, task coordinators and student interns. Each participant was assigned one or more courses and was responsible for reviewing both source-language subtitles and automatically translated versions. The resulting work contributed substantially to increasing the reliability and usability of multilingual educational materials within the Digital Academy.

2. Objectives of the Activity

The activity pursued three complementary objectives:

1. Improving the linguistic quality of automatically generated subtitles.
2. Increasing accessibility and multilingual availability of Digital Academy courses.
3. Integrating Alliance activities into university curricula through recognised internship opportunities.

Students were asked to revise subtitle files, identify transcription and translation errors, improve linguistic quality, verify synchronization with audiovisual content and prepare analytical reports documenting the problems encountered and the corrections introduced. The activity therefore combined quality enhancement, educational innovation and student engagement within a single framework.

3. Organisation of the Activity

The activity was implemented through internship schemes organised by participating institutions.

At the **Università degli Studi di Palermo**, internships were formally integrated into students' study programmes. **Twelve** students participated in the initiative, including ten students from UniPa, one student from JGU and one student participating through inter-university cooperation between UO and UV. Each intern worked under the supervision of the WP3 Task Manager and received continuous academic support. Periodic evaluations and structured feedback sessions ensured the monitoring of progress and alignment with the expected learning outcomes.

At **Universitat de València**, **five** students carried out proofreading activities as part of

internships. Students were provided with subtitle materials by the local IT team and worked under the supervision of academic staff of the Faculty of Philology, Translation and Communication. Their work was subsequently evaluated through individual and collective review sessions.

An additional contribution was made by one PhD student from UO, who developed and tested advanced subtitling methodologies and produced a detailed analytical report concerning subtitle accessibility, technical standards and workflow optimisation. Although the entire assigned course was not completed during the placement period, the methodological contribution represents a valuable resource for future FDA subtitling activities.

4. Methodology

A common methodology was adopted across participating institutions.

Students received subtitle files generated automatically through speech recognition and machine translation systems. Their tasks included:

- revision of transcription errors;
- correction of translation errors;
- verification of linguistic accuracy;
- adaptation of terminology, punctuation and grammar;
- verification of subtitle synchronization;
- reconstruction of omitted passages;
- preparation of analytical reports describing the intervention performed.

Each student completed at least one fully revised subtitle package. The number of courses assigned depended on internship duration and ECTS requirements. Before final submission, participants took part in peer-review discussions in which proofreading strategies, translation choices and workflow challenges were collectively analysed.

The reports reveal that students frequently had to go beyond simple proofreading and perform partial reconstruction of subtitles due to the poor quality of some automatically generated outputs. Recurrent issues included omitted sentences, repeated subtitle segments, invented words, mistranscribed terminology, incorrect timing and untranslated passages.

5. Courses and Languages Covered

The initiative covered courses from a wide range of academic disciplines, including medicine, psychology, statistics, pedagogy, law, communication studies, management, anthropology and sports sciences.

The proofreading activity involved the following language combinations:

Source Language	Target Language
Italian	English
Spanish	English

Source Language	Target Language
English	French
English	German
Spanish	Italian
French	Italian
English	Spanish

At **Jyväskylän Yliopisto** the proofreading of subtitles was carried out directly by the course instructors themselves.

The resulting multilingual output significantly expanded the linguistic accessibility of FDA materials. According to the consolidated data:

- 42 courses were subtitled
- 10 courses were translated into English;
- 2 courses into French;
- 2 courses into German;
- 4 courses into Italian;
- 2 courses into Spanish;
- 1 course from JYU was translated into several languages.

A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2022-23	JYU	Nature under threat 1: Species and Ecosystems	English (orig.)
2023-24	JYU	Children's rights and participation	English (orig.)
2023-24	JYU	Nature under threat 2: Red-List Assessments in Finland and Globally	English (orig.)
2024-25	JYU	Good life and planetary well-being	English (orig.)
2024-25	JYU	Introduction to planetary well-being	English (orig.)
2024-25	JYU	Pathways to planetary well-being	English (orig.)

A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2024-25	JYU	Systems and planetary well-being	English (orig.)
2025-26	JYU	SUS-Waste: Implementing Sustainability in Waste Management	several languages (which ones: list the ones you know)
00 FF	LU	Latvian language for beginners	Yes
00 FF	LU	Latvian language for beginners	Yes
2025-26	LU	Latvian language for beginners	Yes
00 FF	UBE	OWU1 - Open Wine University 1	French (orig.) Italian
A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
00 FF	UBE	OWU2 - Open Wine University 2	French (orig.) Italian
2025-26	uniPa	Analog spectrum analyser and practical measurement techniques	English (orig.)
2025-26	uniPa	Anestesia e Rianimazione	Italian (orig.) English
2025-26	uniPa	ANTROPOLOGIA DEL DESIGN	Italian (orig.) English
2025-26	uniPa	Applied Electronics - Module 1: Linear power supplies. Frequency mixers. PLL.	English (orig.)
2025-26	uniPa	Applied Electronics - Module 2: Analogue modulation basics. Modulators. Demodulators.	English (orig.)

A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2025-26	uniPa	Applied Technology on Mediterranean Food	English (orig.)
2025-26	uniPa	DIGITAL SIGNAL PROCESSING	English (orig.)
2025-26	uniPa	Electronic Instruments and Measurements for Automation	English (orig.)
2025-26	uniPa	European and International Criminal Procedure	English (orig.) French
2025-26	uniPa	Financial Securities, Portfolio and Asset Pricing Models	English (orig.)
2025-26	uniPa	FUNDAMENTALS OF ELECTRONICS	English (orig.)
A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2025-26	uniPa	INNOVATIVE TEACHING APPROACHES AND MOODLE COURSE CREATION	English (orig.) German Spanish
2025-26	uniPa	Introduction to sustainable industrial processes	English (orig.)
2025-26	uniPa	Introduzione ai Modelli Statistici Lineari	Italian (orig.) English
2025-26	uniPa	INTRODUZIONE ALLA PSICOLOGIA GENERALE	Italian (orig.) English
2025-26	uniPa	Medicina di Genere: Disparità, Sfide e Strategie di Comunicazione	Italian (orig.) English
2025-26	uniPa	MICROWAVE AND TERAHERTZ COMMUNICATIONS	English (orig.)

A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2025-26	uniPa	NANOELECTRONICS	English (orig.)
2025-26	uniPa	Nutrition and Supplements in Sports	English (orig.) German French
2025-26	uniPa	Research and Communication Skills – Part 1	English (orig.)
2025-26	uniPa	Research and Communication Skills - Part 2	English (orig.)
2025-26	uniPa	STRATEGIC MANAGEMENT AND OPEN INNOVATION	English (orig.) Spanish
2025-26	uniPa	SUSTAINABILITY MANAGEMENT COURSE	English (orig.)
A.y.	UNIVERSITY	COURSE TITLE	Proofread Sub
2026-2027	uniPa	BASIC ELECTRONICS	English (orig.)
00 FF	UV	Biología molecular: bases y aplicaciones	Spanish (orig.) English
00 FF	UV	Historia de la Ciencia. Introducción General	English
00 FF	UV	Incipit Curso Básico de Lengua y Cultura Latinas	Spanish (orig.) English Italian
00 FF	UV	Didactica y organización socioeducativa	Spanish (orig.) English Italian
2022-23	UV	History of Science. A general introduction	English

6. Quantitative Results

The activity produced measurable outputs that exceeded the original target of proofreading fifteen FDA courses.

The following results were achieved:

- 42 original subtitle transcriptions proofread;
- 21 translated subtitle versions proofread;
- 10 curricular internships formally recognised with ECTS credits at UniPa;
- 6 curricular internships formally recognised with ECTS credits at UV.

Overall, 18 students participated directly in proofreading activities through internships and some teachers provide subtitles to their own courses. The initiative generated not only revised subtitle files but also a substantial collection of analytical reports documenting recurring errors, best practices and recommendations for future workflow improvements.

7. Main Linguistic and Technical Findings

A comparative analysis of student reports reveals highly consistent patterns across different courses and language combinations.

The most frequently encountered issues concerned automatic speech recognition errors. Proper names, specialist terminology and multilingual expressions were often mistranscribed. Students reported numerous cases in which technical vocabulary from medicine, law, psychology, pedagogy and statistics was either incorrectly recognised or replaced with unrelated words.

Translation quality also varied considerably. Literal translations, incorrect terminology choices, false friends and semantic distortions were common. In some cases, translations radically altered the intended meaning of the source content. This was particularly evident in specialised courses where precise terminology plays a crucial role, such as medicine, criminal procedure and educational sciences.

Technical issues represented another major challenge. Students repeatedly documented:

- duplicated subtitle blocks;
- looped sentences;
- omitted subtitle segments;
- synchronization problems;
- incorrect segmentation;
- timing inconsistencies;
- generation of non-existent words;
- appearance of random characters from unrelated writing systems;
- untranslated portions of text in target-language files.

Several reports concluded that the amount of human intervention required often exceeded expectations and transformed proofreading activities into comprehensive subtitle reconstruction work. Particularly severe problems were observed in medical and legal courses, where entire sections occasionally required manual recreation.

Despite these challenges, students also identified positive aspects. In many cases the automated tools provided a useful starting point, reducing the amount of manual transcription required. The overall experience demonstrated that AI-generated subtitles can effectively support the workflow, provided that systematic human revision remains an integral stage of the process.

8. Student Feedback and Learning Outcomes

Student reports consistently describe the activity as a valuable and highly formative learning experience. Beyond the immediate proofreading tasks, participants highlighted the opportunity to apply theoretical knowledge acquired during their degree programmes to authentic multilingual educational materials. The activity enabled students to **engage directly with real-world translation and accessibility challenges** while developing practical **professional skills relevant to future careers in translation, communication, education and international cooperation**.

One of the most recurring themes concerns the acquisition of **advanced revision and post-editing competences**. Students learned to compare audio recordings with generated subtitles, identify inconsistencies, reconstruct omitted content and evaluate the reliability of machine-generated translations. Many participants reported that the activity strengthened their analytical abilities and their awareness of linguistic nuances, particularly when dealing with multilingual and specialised academic content. Students working on courses in medicine, law, pedagogy, psychology, management and anthropology frequently emphasised the importance of contextual analysis for selecting appropriate terminology and maintaining semantic accuracy.

Another important learning outcome concerned **terminology management**. Participants repeatedly noted that the automated systems encountered significant difficulties when processing specialised vocabulary. As a consequence, students had to familiarise themselves with discipline-specific terminology and independently verify meanings through research and documentation. This process contributed to the development of specialist language competences and a greater understanding of subject-specific discourse. Students observed that terminology-related errors represented some of the most consequential problems because they directly affected the educational value of the courses and, if left uncorrected, could lead to misunderstandings among learners.

Several reports highlighted the **educational value** of working with multiple languages simultaneously. Students revised materials involving Italian, English, French, German and Spanish, often moving between source texts, translated subtitles and audiovisual content. This multilingual environment strengthened intercultural awareness and reinforced their ability to operate within international and collaborative contexts. Students reported improvements in translation skills, linguistic accuracy and cultural adaptation strategies, all of which are highly relevant in contemporary multilingual educational settings.

Many participants reflected extensively on the **strengths and limitations of artificial intelligence**. While acknowledging that automated systems can accelerate initial transcription and translation processes, students concluded that current AI-generated subtitles cannot be considered reliable without comprehensive human revision. Reports repeatedly describe grammatical mistakes, lexical inaccuracies, mistranslations, omissions, synchronization problems and even completely fabricated content generated by automated tools. Some students concluded that correcting poor-quality machine output occasionally required more effort than creating subtitles manually from the original audio. These observations provided participants with a realistic understanding of AI-assisted workflows and the continuing importance of human expertise in quality assurance processes.

The internships also contributed to the **development of transferable skills**. Students frequently mentioned improvements in time management, problem solving, attention to detail, independent decision-making and collaborative work. The requirement to produce analytical reports encouraged reflective practice and allowed students to critically evaluate both their own work and the technologies employed during the project. Feedback sessions organised at the end of the internship period further supported this process by enabling participants to compare experiences, discuss effective revision strategies and collectively examine the impact of AI tools on multilingual educational resources.

Overall, student feedback demonstrates a **high level of satisfaction** with the activity. Participants considered the internships particularly valuable because they closely matched their academic backgrounds and offered a meaningful opportunity to apply theoretical knowledge in a professional context. Many students emphasised that the experience provided a clearer understanding of accessibility, audiovisual translation and multilingual communication while simultaneously reinforcing the practical relevance of their university studies. The activity therefore achieved not only its technical objectives but also a significant educational impact on participant learning and professional development.

9. Impact on Accessibility and Multilingualism

The activity has made a direct contribution to the strategic objectives of the FORTHEM Alliance regarding **inclusiveness, multilingualism and educational accessibility**.

The revision and correction of subtitles significantly improved the quality, reliability and usability of Digital Academy content. Students corrected errors capable of compromising both comprehension and accessibility, thereby ensuring that learners from different linguistic backgrounds can access the educational materials more effectively.

A particularly relevant outcome concerns the expansion of language coverage. Beyond English, several courses originally delivered in Spanish have already been translated and proofread in Italian, thereby reaching additional student groups and further enhancing accessibility across partner institutions.

10. Integration into Curricula and Good Practice

One of the most significant achievements of the activity was the successful integration of proofreading and subtitling work into **recognised curricular internships**.

The allocation of ECTS credits transformed the activity from a voluntary contribution into a structured learning experience formally embedded within university programmes. Students received academic recognition while simultaneously contributing to a strategic Alliance initiative. This approach created tangible benefits for both institutions and learners.

The activity therefore represents a particularly successful example of how European University Alliance initiatives can be integrated into ordinary university curricula while generating high-impact outputs that serve both institutional objectives and student development.

11. Lessons Learned and Recommendations

The reports produced by students provide a number of important recommendations for future activities.

First, **AI-generated subtitles** should always undergo systematic human review before publication. Automated transcription and translation technologies may provide an efficient starting point, but current outputs remain insufficiently reliable for direct educational use. Second, **common subtitling guidelines** should be adopted across the Alliance to improve consistency. Recommendations emerging from student reports include clearer standards for timing, segmentation, readability and accessibility. Particular attention should be devoted to subtitle synchronization and compliance with established international subtitling conventions. Third, the **internship model** should be maintained and expanded. The positive educational outcomes, the production of high-quality multilingual resources and the successful integration of ECTS recognition demonstrate the effectiveness of this approach.

The **contribution of the DAO IT staff** at the FPU (specifically at UBE, UV and uniPa) proved to be essential, both for providing subtitles and uploading them. The limited technical expertise of some of the students involved resulted in subtitles that were technically imperfect, leading to issues related to synchronization and subtitle length. At this stage, to address the issues efficiently and consistently, the uniPa DAO IT team developed a **dedicated Python script** that automatically corrected synchronization and formatting problems, significantly improving the quality and usability of the subtitles.

12. Conclusion

This activity successfully exceeded its original objective of proofreading fifteen FDA courses and produced substantial educational, technical and institutional benefits. Through the participation of students from multiple FORTHEM universities, a large number of subtitle files and translated versions were revised, corrected and improved, contributing directly to the accessibility and multilingual reach of the FORTHEM Digital Academy.

The initiative demonstrated that although AI-assisted subtitling technologies can support multilingual content production, human expertise remains essential for ensuring linguistic accuracy, technical quality and educational reliability. At the same time, the activity provided students with valuable professional experience and successfully integrated Alliance objectives into recognised academic pathways through internship schemes and ECTS accreditation.

The results achieved confirm that this model constitutes a highly effective good practice for future FORTHEM activities aimed at promoting accessibility, multilingualism, student engagement and the sustainable development of shared educational resources.

JOHANNES GUTENBERG-UNIVERSITÄT MAINZ

Saarstraße 21,
55122 Mainz, Deutschland

**JYVÄSKYLÄN YLIOPISTO**

Seminaarinkatu 15
40014 Jyväskylän yliopisto, Suomi

**LATVIJAS UNIVERSITĀTE**

Raiņa Blvd.19,
Rīga, 1586, Latvija

**UNIVERSITÀ DEGLI STUDI DI PALERMO**

Piazza Marina, 61
90133 Palermo, Italia

**UNIVERSITAT DE VALÈNCIA**

Avda. Blasco Ibañez, 13
46010 València, España

**UNIVERSITATEA LUCIAN BLAGA DIN SIBIU**

Bulevardul Victoriei 10,
Sibiu 550024, România

**UNIVERSITÉ DE BOURGOGNE**

Maison de l'Université, Esplanade Erasme
21078 Dijon Cedex, France

**UNIVERSITETET I AGDER**

Universitetsveien 25,
4604 Kristiansand, Norge

**UNIwersYTET OPOLSKI**

Pl. Kopernika 11a
45-040 Opole, Polska

