

Subtitling video lectures using Miro Translate

Laura Cacheiro Quintas¹, Samuel Calegari².

¹University of Perpignan Via Domitia, 52, avenue Paul Alduy, 66100, laura.cacheiro@univ-perp.fr

²University of Perpignan Via Domitia, 52, avenue Paul Alduy, 66100, samuel.calegari@univ-perp.fr

Keywords

Subtitling, captions, automatic speech recognition, neural machine translation, video lectures, e-Learning.

1. Summary

Miro Translate is a cloud-based platform that produces captions and subtitles for video lectures. This hybrid solution uses an Automatic Speech Recognition (ASR) system to produce machine-generated captions and Neural Machine Translation (NMT) to elaborate machine-generated subtitles. Users have at their disposal a set of parameters and editing functionalities to adapt the target text to different quality standards and to the Deaf and Hard of Hearing regulations in force in Europe. Therefore, audiovisual content is accessible to non-native speakers and to people suffering from hearing loss or deafness.

2. BACKGROUND

The number of students enrolled in distance learning is increasing, including learners with disabilities. In this regard, most European countries have ratified the Convention on the Rights of Persons with Disabilities (2006) and have adopted domestic regulations in favor of accessibility. Video lectures displaying captions and subtitles become a cost-efficient solution to communicate complex messages in an attractive manner because they integrate written text, images and speech through visual and auditory channels (Díaz Cintas, 2014). The growing demand for captions and subtitles in higher education requires specific tools. When designing the software architecture two main parameters must be considered. First, subtitling conventions and explicit extratextual norms require specific editing options for their implementation. Since these regulations reflect the target audience expectations and are cultural-dependant, the software must present a certain degree of adaptability. Second, teachers, clerical or technical staff usually produce captions and subtitles without specific training. For this reason, an intuitive interface is required.

3. A HYBRID SUBTITLING TOOL

Miro Translate is part of an ongoing eLearning research project conducted by the MIRO.EUPM Programme, which is part of the French National Research Institute (ANR) Initiatives for Excellence in Innovative Training (IDEFI). It is a hybrid cloud-based subtitling platform to produce captions and subtitles of instructional videos. Its functionalities allow for different degrees of automation (manual, automatic or hybrid mode) depending on the level of quality required and the time available.

When using the manual mode, users upload a video to the platform and manually perform the text input and the spotting (i.e. synchronization of the subtitles with the soundtrack). The program proposes various tools and settings according to the linguistic and technical conventions such as reading speed parameters, typeface or speaker's identification options. However, this is a time consuming task.

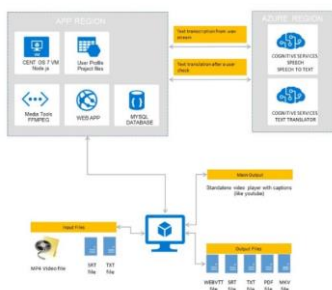


Figure 1. System Description

When using the automatic mode, the user will login to the Miro Translate platform, and send the audio to the Microsoft Services Speech, a speech-to-text system that generates automatic captions. The output is displayed in the platform and subtitlers choose whether to edit the auto-generated transcriptions using the functionalities available or to use them without any modification.

For subtitling, users open an existing project and send text information to the Microsoft Text Translator, a text-to-text system that performs neural machine translation. The automatic-generated subtitles are displayed in the platform for post-editing when necessary. Different output files are available supporting metadata related to time-alignment, line breaks, font style and position. Sound track information can also be exported as a text file allowing students to manipulate the target text and use it as notes. Finally, when exporting an MKV file, videos can be watched offline on any support with different subtitles tracks.

4. CONCLUSIONS

Manually transcribing and translating the fast-growing number of video lectures using subtitling programs is expensive and time-consuming. This situation calls for new solutions that find synergies between artificial intelligence and human subtitlers. Auto-generated captions and subtitles offer a solution to speed up text input and time-alignment. However, automatic-generated outputs need to be improved to comply with subtitling quality standards and conventions. Miro Translate is a tool adapted to eLearning environments developed in respond to this challenge. This work is part of an ongoing study and further efficiency and usability tests will be performed with the intention of improving system performance.

5. REFERENCES

AENOR website (2012). *Norma UNE 153010. Subtitulado para personas sordas y personas con discapacidad auditiva*. Retrieved February 8, 2019, from http://implantecoclear.org/documentos/accesibilidad/UNE_153010_2012.pdf

Díaz Cintas, J. (2014). *Technological Strides in Subtitling*. In Chan, S-W, editor in chief, Routledge Encyclopedia of Translation Technology. Routledge, London, pages 632-643.

Ford Williams, G. (2009). *Online Subtitling Editorial Guidelines*. Retrieved February 8, 2019, from http://www.bbc.co.uk/guidelines/futuremedia/accessibility/subtitling_guides/online_sub_editorial_guidelines_vs1_1.pdf

Ivarsson, J., and Carroll, M. (1998). *Code of good subtitling practice*. In Language Today, April: 1-3. <https://www.esist.org/wp-content/uploads/2016/06/Code-of-Good-Subtitling-Practice.PDF.pdf>

Karamitroglou, F. (1998). *A Proposed Set of Subtitling Standards in Europe*. In Translation Journal, 2:2. Retrieved February 8, 2019, from <http://translationjournal.net/journal/04stndrd.htm>