

Creating and translating audio descriptions with the help of *Frazier*



Eva Schaeffer-Lacroix (MCF HDR)

Sorbonne Université, STIH, section E3 (Computer linguistics)

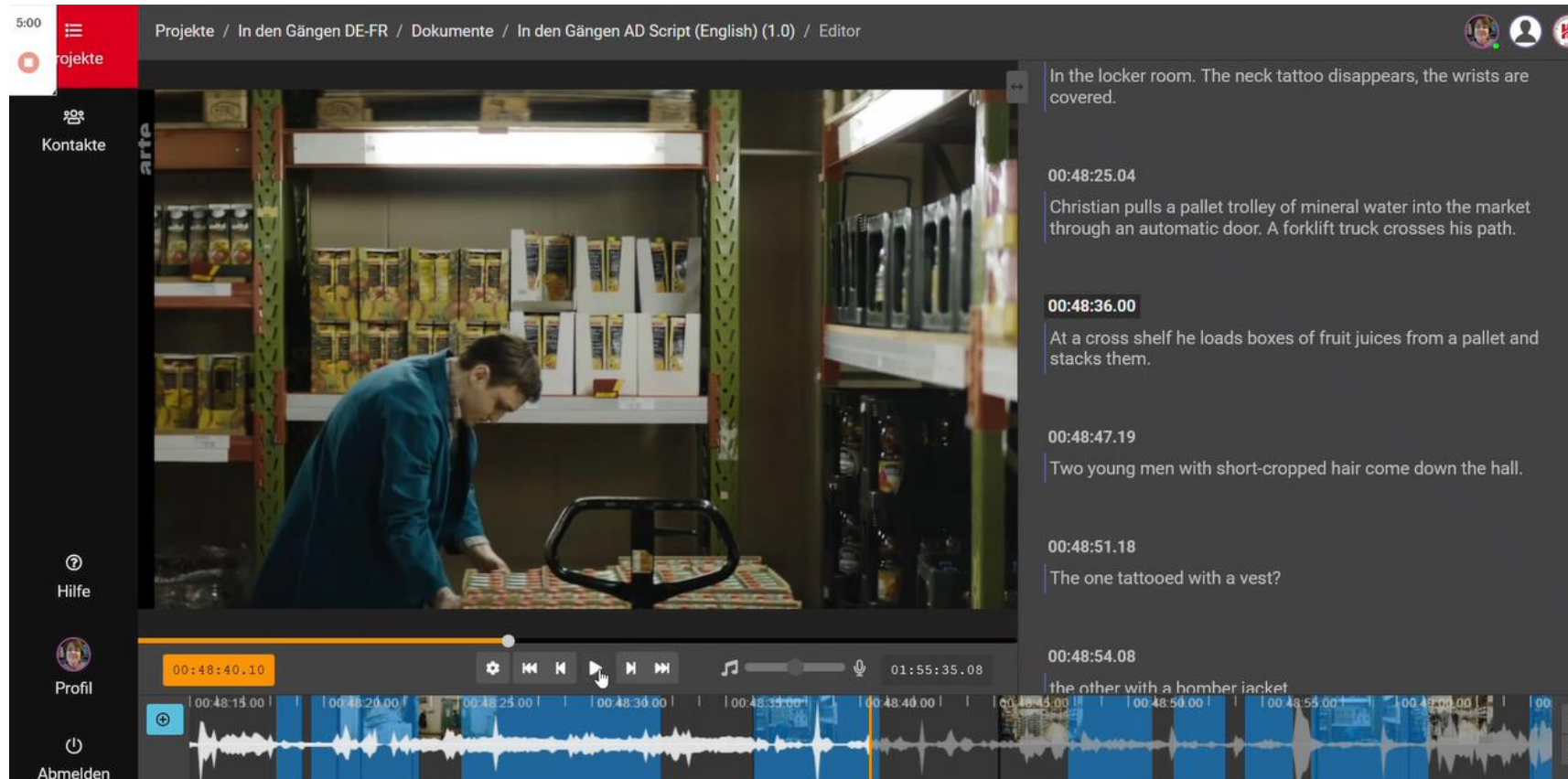
Saskia Josephine Schulz (Research Assistant)

Institute for Translation Studies and Specialised Communication University of Hildesheim

Outline

- Introduction into audio description
- TADS project (Translation of Audio Description Scripts)
- *Frazier*
 - Functions
 - How to create audio descriptions
 - How to translate audio descriptions (& DeepL)
 - How to use the text-to-speech function
- Comparison of the AD obtained by machine translation with the author ADs
- Tool evaluation
- Implications for research

Audio description sample: *In den Gängen*



The screenshot shows a video player interface for the film 'In den Gängen'. The video frame displays a man in a blue jacket working in a market aisle, loading boxes of fruit juices onto a pallet. The interface includes a top navigation bar with the text 'Projekte / In den Gängen DE-FR / Dokumente / In den Gängen AD Script (English) (1.0) / Editor'. A left sidebar contains navigation options: 'Kontakte', 'Hilfe', 'Profil', and 'Abmelden'. The video player controls at the bottom show a progress bar at 00:48:40.10 and a volume icon. An audio description overlay on the right side of the video frame provides the following text:

In the locker room. The neck tattoo disappears, the wrists are covered.

00:48:25.04
Christian pulls a pallet trolley of mineral water into the market through an automatic door. A forklift truck crosses his path.

00:48:36.00
At a cross shelf he loads boxes of fruit juices from a pallet and stacks them.

00:48:47.19
Two young men with short-cropped hair come down the hall.

00:48:51.18
The one tattooed with a vest?

00:48:54.08
the other with a bomber jacket

An audio description is a sound track which is added to a film in order to help blind or visually impaired people understand what happens in the film.

Audio description script

10:50:53

s Der Tätowierte hält zwei Schnapsflaschen. Sein Einkaufswagen ist voller Getränke.

10:51:01

„*Arbeitest du hier? - Nee!*“
(*bisschen Lachen übersprechen*)
Christian schaut sich verlegen um und geht auf die beiden zu.

s (schnell)

10:50:53

f The tattooed man is holding two liquor bottles. His shopping cart is full of drinks.

10:51:01

„*You are working here? - Nah!*“
(*crosstalk a bit of laughter*)
Christian looks around sheepishly and walks up to the two.

f (fast)

Brief research overview

- Around 2007, research on audio description (AD) starts in many disciplines (e.g. in applied linguistics).
- Research focus lies on the AD of artefacts like films and museum exhibitions.
- The language pairs "German-French/French-German" are not well represented in translation studies focussing on AD.

The TADS project

- **Origin:**

Initiated by a French-German research team (Sorbonne Université & Universität Hildesheim)

- **Research question:**

How to best produce AD scripts for the dubbed version of movies, TV shows, and documentaries?

- **Working assumption:**

For some target languages, translating AD manuscripts seems to be a good alternative to direct AD manuscript production (Jankowska et al. 2015; Jankowska et al. 2017), called "author ADs".

- **Choices to be made:**

Human translation (HT) and/or machine translation (MT) with post-editing?

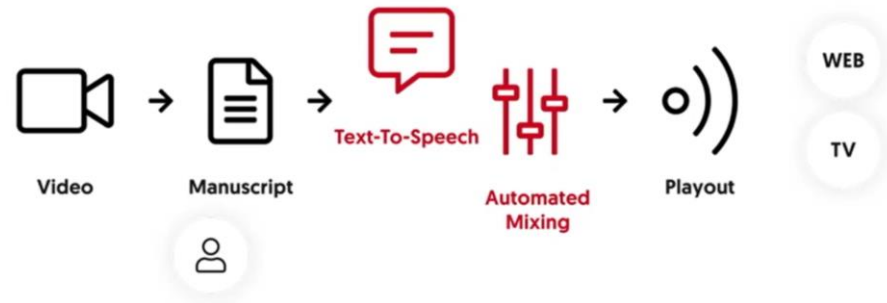
Hypotheses

- **Hypothesis 1:** Since AD scripts contain strongly standardised recurrent elements (Kleege 2016), machine translation is likely to lead to good results at least for translating certain parts of ADs (Fernández-Torné 2016).
- **Hypothesis 2:** Localisation tasks (cultural adaptation and user orientation) are better handled by human translators than by MT.

Project tools

- The corpus management system TXM to do textometric analysis (Heiden et al. 2010)
- The transcription tool EXMARaLDA to display the videos and the ADs of the original and of those produced for other languages (Schmidt 2012)
- *Frazier* to create ADs and/or to automatically translate ADs

Frazier functions



- Manual insertion of parts to be recorded & automatic creation of time slots
- Synthetic voice (text-to-speech)
- Automatic volume mixing
- Automatic translation of added audio descriptions
- Various export options: srt, vtt, csv, .docx, .adc

How to create an AD with *Frazier*

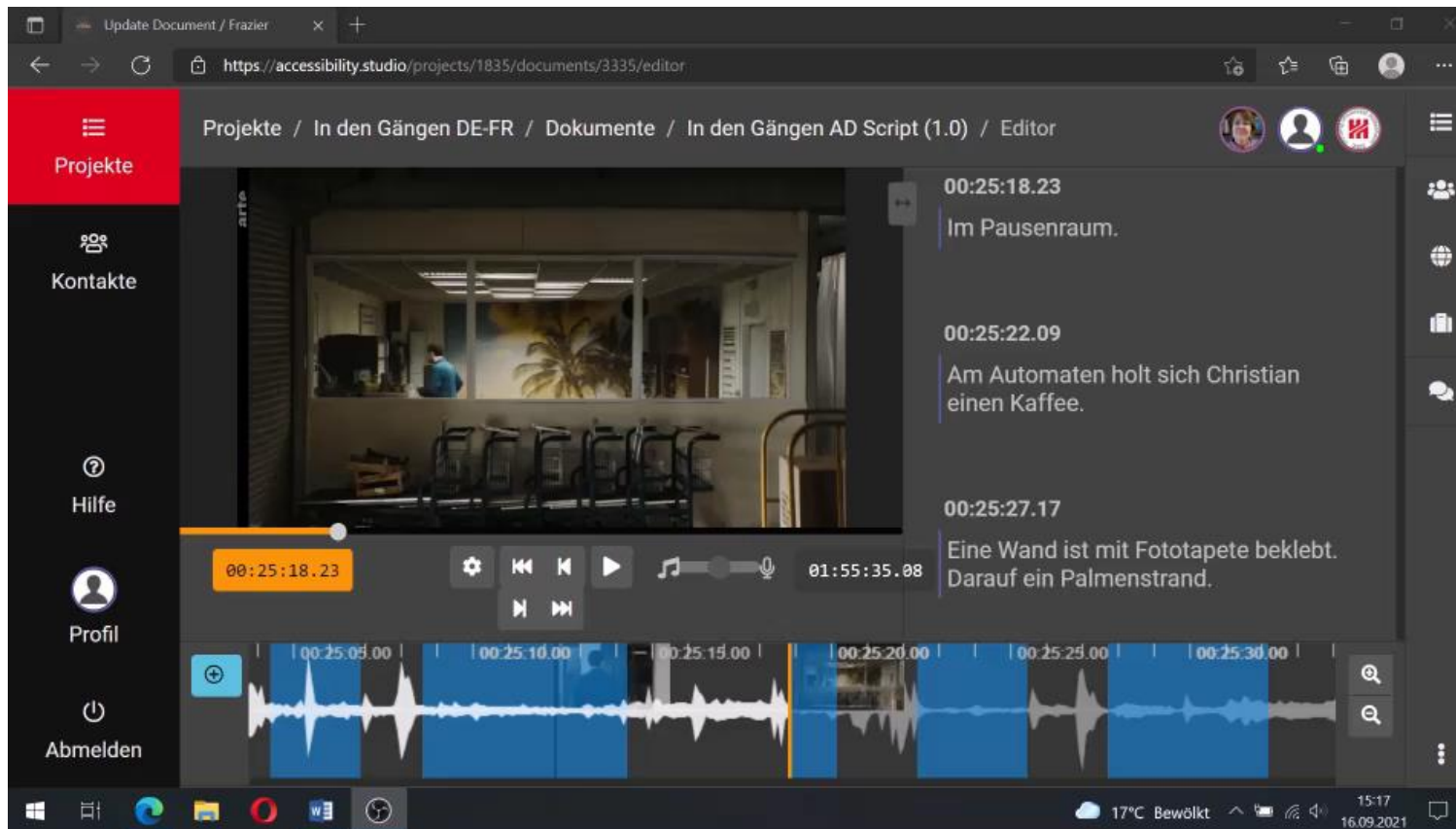
Preparation

- To produce an AD with *Frazier*, one must start by creating the project, selecting the language, and uploading the video file. In case your video file is already equipped with a timecode, there is also an option to adjust it. Once the script file is generated, it is all pretty intuitive. By clicking on the script file, you enter the insertion mode, where you can watch and listen to the video file and jump to the frame of your choice.

AD text passages

- In the lower left corner, there is a blue button with a plus on it, which opens a slot for an AD text passage. Since *Frazier* is a text-to-speech tool, it generates a spoken version of the text passage, read by an artificial voice. The length of the passage is calculated automatically, but you can slightly adjust it by selecting the individual speed of the passage and even the voice reading it. Normally, the artificial voice is chosen at the beginning of the project, but can be changed for every single passage if necessary.

How to create an AD with *Frazier*



Once the AD script passage is inserted, you can watch the video with the AD to check whether the timing, length, and wording are all to your liking.

How to create an AD with *Frazier*

Teamwork options

- The administrator of the project can add other team members to work simultaneously on the AD. Thanks to the chat function, the team members can comment on certain passages using timestamps.
- Furthermore, *Frazier* allows to create a sort of terminology database for characters, places, and objects. All the changes they make, AD script or comments, are saved automatically on the server so that they also show up immediately to the rest of the team.

Ducking function

- Automated mix between AD and the audio track of the video
- Different broadcasting standards
 - between countries
 - between TV and online media



The screenshot shows a dark-themed interface for an 'Automated Mix' tool. It features several dropdown menus and a list of options. The 'Mix Type' dropdown is set to 'No ducking - keep original audio track untouched'. The 'Mastering / Loudness Target' dropdown is open, showing a list of options: '-13 LUFS (very loud, max peak 0db)', '-16 LUFS (Web, PodCasts, Mobile)', '-23 LUFS, EBU R 128 (TV, Europe)', and '-24 LUFS, ATSC A/85 (TV, US)'. The 'Original Audio Track' dropdown is set to 'big-buck-bunny.wav'. The 'Voice' dropdown is set to 'big-buck-bunny-audio-description_tts_202102131404.wav'. At the bottom, there are 'Back' and 'Start' buttons.

<https://www.videotovoice.com/knowledge-base/mixing-mastering/>

How to use *Frazier* to translate ADs

Translating an AD is a fairly easy task once you've discovered the three little dots in the lower right corner. Simply click on the option "translate document" and wait a moment. However, if you plan to keep the original source-language text, you need to duplicate the script document first and have the copy translated. The translation in Frazier is fully automated: you just need to select the target language and the voice, and then it generates all the passages in the target language.

I've had the translation done into English. This makes it possible to view the result with an English AD. It might need some minor post-editing, like changing a word or two or moving the passage around if it has become significantly longer or shorter than the original, but these are minor changes like those needed when using other machine translation tools.

Limits of the result obtained by MT (DeepL)

German author AD

10:12:37

Bruno grinst nachsichtig.

10:12:43

Bedächtig zieht Christian den Wagen mit drei leeren Paletten in den breiten Hauptgang.

10:13:21

Neugierig betrachtet er die Deichsel.

French MT AD

10:12:37

Bruno sourit avec indulgence.

→ Length

10:12:43

Christian tire pensivement le chariot avec trois palettes vides dans la grande allée principale.

→ Erroneous syntax

10:13:21

Curieusement, il regarde le timon.

→ Erroneous meaning

Limits of the result obtained by MT (DeepL)

German author AD

10:12:37

Bruno grinst nachsichtig.
[Bruno grins indulgently.]

10:12:43

Bedächtig zieht Christian den Wagen mit drei
leeren Paletten in den breiten Hauptgang.
[Christian carefully pulls the trolley, loaded with
three empty pallets, into the wide main aisle.]

10:13:21

Neugierig betrachtet er die Deichsel.
[He casts an interested glance at the drawbar.]

French MT AD

10:12:37

Bruno sourit avec indulgence.
→ Length

10:12:43

Christian tire pensivement le chariot avec trois
palettes vides dans la grande allée principale.
→ Syntax and meaning

10:13:21

Curieusement, il regarde le timon.
[Strangely, he looks at the drawbar.]
→ Meaning

French MT AD versus author ADs

German author AD	French MT AD (DeepL)	French author AD
<p>10:50:53 Der Tätowierte hält zwei Schnapsflaschen. Sein Einkaufswagen ist voller Getränke. [The tattooed man is holding two liquor bottles. His shopping trolley is full of drinks.]</p>	<p>48:56 L'homme tatoué tient deux bouteilles d'alcool. Son caddie est rempli de boissons.</p>	<p>49.31 Le jeune homme, blond, tatoué dans le cou, s'approche et affronte Christian du regard. [The young man, blond, tattooed on his neck, approaches and challenges Christian with his eyes.]</p>
<p>10:51:01 Christian schaut sich verlegen um und geht auf die beiden zu. [Christian looks around sheepishly and walks towards the two.]</p>	<p>49:04 Christian regarde autour de lui d'un air penaud et se dirige vers les deux.</p>	<p>49.40 Christian fait oui et s'éloigne. [Christian nods and goes away.]</p>

Tool evaluation

Frazier is a powerful audio description tool:

- It allows perfect and comfortable timeslot management.
- Its text-to-speech function provides very natural sounding speech.
- Machine translation works fine.
- It offers various helpful teamwork options.

Implications for AD research

Which are the limits of automatic AD creation, translation, and voicing?

- Users may not appreciate synthetic voices, depending on the sort of film.
- Automatic pronunciation and prosody needs further training.
- Machine translation needs further specific training with AD data.

Is it possible to reduce the huge gap between French and German author ADs?

- No: Cultural and linguistic differences justify the use of the author AD creation method.
- Yes: Common AD production methods could be developed.

Acknowledgements

- Sommerhaus Produktion & permission to use the video *In den Gängen* for research purposes
- ARTE & permission to use the French and German AD scripts of this movie
- Antonia Lins & comparison of translation tools
- Christian David and Lukas Pajoncsek, Video to voice & permission to use of *Frazier* for the needs of the TADS project

References

- Fernández-Torné, Anna. 2016. 'Machine Translation Evaluation through Post-Editing Measures in Audio Description'. *InTRAlinea* 18. <http://www.intraline.org/archive/article/2200>.
- Heiden, Serge, Jean-Philippe Magué, and Bénédicte Pincemin. 2010. 'TXM : Une plateforme logicielle open-source pour la textométrie - conception et développement'. In *Proceedings of 10th International Conference Journées d'Analyse statistique des Données Textuelles*, 2(3):1021–32. Edizioni Universitarie di Lettere Economia Diritto. <https://halshs.archives-ouvertes.fr/halshs-00549779/document>.
- Jankowska, Anna, Anna Mrzyglodzka, and Anna Chociej. 2015. *Translating Audio Description Scripts: Translation as a New Strategy of Creating Audio Description*. Text - Meaning - Context: Cracow Studies in English Language, Literature and Culture, volume 12. Frankfurt am Main ; New York: Peter Lang Edition.
- Jankowska, Anna, Michał Milc, and Louise Fryer. 2017. 'Translating Audio Description Scripts... into English'. *THE SLOVAK ASSOCIATION FOR THE STUDY OF ENGLISH*, 2017, sec. 10(2).
- Kleege, Georgina. 2016. 'Audio Description Described: Current Standards, Future Innovations, Larger Implications'. *Representations* 135 (1): 89–101. <https://doi.org/10.1525/rep.2016.135.1.89>.
- Pajoncsek, Lukas, and Christian David. 2019. Frazier. VIDEO TO VOICE GmbH. <https://www.videotovoice.com/>
- Reviere, Nina. 2018. 'Studying the Language of Dutch Audio Description: An Example of a Corpus-Based Analysis'. *Translation and Translanguaging in Multilingual Contexts* 4 (1): 178–202. <https://doi.org/10.1075/ttmc.00009.rev>.
- Schmidt, Thomas. 2012. 'EXMARaLDA and the FOLK Tools – Two Toolsets for Transcribing and Annotating Spoken Language'. In *Proceedings of the 8th International Conference on Language Resources and Evaluation: May 23-25, 2012*, edited by Thierry Declerck, Khalid Choukri, and Nicoletta Calzolari, 236–40. Istanbul (Turquie).
- Zago, Raffaele. 2019. 'English Films vs Italian Films: A Comparative Analysis via the Pavia Corpus of Film Dialogue and the WordSmith Tools'. In *Benjamins Translation Library*, edited by Irene Ranzato and Serenella Zanotti, 148:230–41. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/btl.148.11zag>.

Contact

Eva Schaeffer-Lacroix

- Email: evalacroix@free.fr
- Orcid : <https://orcid.org/0000-0002-6260-9095>
- Academic homepage: <http://didaktik.hautetfort.com>

Saskia Josephine Schulz

- Email: schul026@uni-hildesheim.de