10th ADVANCED RESEARCH SEMINAR ON AUDIO DESCRIPTION

TransMedia Catalonia Research Group

Universitat Autònoma de Barcelona

Venue: Institut d'Estudis Catalans

(Carrer del Carme, 47, Barcelona)

19-20-21 March 2025

TABLE OF CONTENT

COMMITTEES, ORGANISERS AND SPONSORS

ARSAD director: Anna Matamala, UAB.

Scientific committee

- Brett Oppegaard, University of Hawai'i at Manoa.
- Elisa Perego, University of Pavia.
- Estel·la Oncins, Universitat Autònoma de Barcelona.
- Gert Vercauteren, University of Antwerp.
- Haina, Jin, Communication University of China.
- Iwona Mazur, Adam Mickiewicz University.
- Lucinea Marcelino Villela, Sao Paulo State University.
- Marie Campbell, Red Bee Media.
- Marina Ramos Caro, University of Murcia.
- Monika Zabrocka, Jagiellonian University.
- Sarah McDonagh, Universitat Autònoma de Barcelona.
- Veronika Rot, RTV Slovenija.

Organising committee (TransMedia Catalonia, UAB)

- Anna Matamala (coordinator).
- Carme Mangiron.
- Chiara Gunella.
- Estel·la Oncins.
- Marina Pujadas.
- Marta Brescia.
- Miguel Ángel Oliva.
- Pilar Orero.
- Sarah McDonagh.

Gold sponsors

Silver sponsors

This conference is part of TransMedia Catalonia research group activities (2021SGR00077), and more specifically of the WEL project (PID2022-137058NB-I00).

ARSAD is also part of AccessCat activities. AccessCat is the Catalan Accessibility Network funded by the Department of Research of the Catalan Government (2021XARDI0007).

Accessibility sponsors

Media partners

CONFERENCE PROGRAMME UPDATED 13.03.2025

#arsad2025
https://webs.uab.cat/arsad

PRE-CONFERENCE WORKSHOP: 19 MARCH 2025

08:30-09:30	Registration
10:45-11:30	Registration
09:00-11:00	**ENACT pre-conference workshop** on media accessibility: state of the art and future challenges *(free registration for ARSAD participants)* - Welcome and overview of ENACT (**Veronika Rot**, RTV Slovenja, ENACT project coordinator) - Audio description and audio subtitling (**Elisa Perego**, University of Pavia) - Subtitling (**Carlo Eugeni**, University of Leeds) - Sign language interpreting (**Marta Bosch**, Direcció General de Política Lingüística, Generalitat de Catalunya) - Easy-to-understand language (**Anna Matamala**, Universitat Autònoma de Barcelona) - Accessibility technologies in public media (radio & television) (**Andrej Tomažin**, RTV Slovenija) - Web and digital accessibility (**Carlos Duarte**, Universidade de Lisboa) - Artificial Intelligence and accessibility (**Sabine Braun**, Surrey University) - Accessibility in virtual reality and virtual worlds (**Pilar Orero**, UAB) - User perspective (**Daniel Casas**, European Disability Forum)

Coffee break (11:00-11:30)

CONFERENCE- DAY 1: 19 MARCH 2025

11:30-11:45	Welcome address by **Anna Matamala** (TransMedia Catalonia, UAB) including short TMC, WEL and AccessCat presentation.
11:45-12:30	**KEYNOTE LECTURE** by **Sabine Braun** (Surrey University). Bridging the Accessibility Gap? The role of AI in the future of audio description generation. Chair: Anna Matamala (UAB). 30' minute presentation plus 15' discussion.
12:30-13:30	**PANEL 1. TECHNOLOGY: INDUSTRY PERSPECTIVES** Chair: Mereijn van der Heijden (Earcatch) - **Joel Snyder** (Audio Description Associates). *AI and AD: For Good or Ill?* (15') - **Matthew Kaplowitz,** Wendy Sapp (Bridge Multimedia). *Additive Prompting: A Tool for Ai Assisted Audio Description* (15') - **Eveline Ferwerda** (Scribit.Pro). *Using AI for Audio Description generation* (10') Discussion: 20 minutes

Lunch break (13:30-14:30) *Although lunch is not included, sandwiches will be offered to participants*

14:30-16:15	**PANEL 2. TECHNOLOGY: RESEARCH PERSPECTIVES** Chair: Estel·la Oncins (UAB) - **Brett Oppegaard** (University of Hawaii at Manoa). *A Strange New World. Behind the Scenes of Designing Novel Gen-AI Approaches to Audio Description* (15') - **Marina Ramos,** Purificación Meseguer (Universidad de Murcia). *Emotions offstage: diversity and violence described by AI?* (15') - **Nicole Lüthi,** Alexa Lintner, Lukas Fischer, Martin Kappus, Sarah Ebling (Zurich University of Applied Sciences). *An exploratory analysis of LLM-based machine-translated audio description scripts in the French-German language pair* (15') - **Blanca Serrano Alcaraz,** Ana María Rojo López, Marina Ramos, Konrad Rudnicki (Universidad de Murcia). *Exploring the Impact AI Voices in Erotic Film Audio Description: A Gender-Based Study* (15') - **Xiaochun Zhang,** Ben Ackland, Thomas Keane (University College London). *AIAD Project: Making Video Games Accessible for the Visually Impaired with Dynamic AI Audio Description* (15') Discussion: 30 minutes

Coffee break (16:15-16:45)

16:45–17:45	**MOSAIC project workshop.** AI meets Audio Description. Exploring opportunities and risks. Organiser: Pilar Orero (UAB).

Social event for panel speakers organised by Bridge Multimedia.

Event by invitation only, details to follow to panel speakers (19:00)

CONFERENCE- DAY 2: 20 MARCH 2025

08:45–09:00	Registration
09:00–10:45	**PANEL 3. AD PROCESSES: looking into the future** Chair: Marina Ramos (Universidad de Murcia) - **Olga Davis,** Dimitris Asimakoulas, Sabine Braun (Surrey University). *Modular Audio Description: A Novel Approach to AD Creation and Delivery* (15') - **Mariana J. López,** Gavin Kearney (University of York). *Enhanced Audio Description (EAD): Making Documentaries Accessible Through Sound Design* (15') - **Marie Campbell** (Red Bee Media). *Storytelling: AD in the Past, Present and Future* (15') - **Tamara Domínguez Ebitsch** (Johannes Gutenberg University Mainz). *Audio Description Quality Standards, 20 years on: Insights from expert interviews with German-speaking B/VIP users* (15') - **Elena SV Flys,** Antonio Villalba (TAI University Center for the Arts). *Live Audio Description During a Film Shooting* (15') Discussion: 30 minutes
10:45–11:15	**AD PITCHES I** (5 minutes each) Chair: Anna Matamala (UAB) - **Sarah Barron**, Janice Rieger (Griffith University). *Co-designing artwork audio description in Australia.* - **Ana-Julia Perrotti-Garcia** (Pontifical Catholic University of Minas Gerais). *Audio description of one of the world's largest street parades: translating Brazilian Carnaval into words.* - **Vanessa Gumier** (Universitat Jaume I). *The voice of the Valencian* ninots*: A reflection on neutrality.* - **José-Luis Castillo-Flores** (Universidad de las Palmas de Gran Canaria). *Testing Clipchamp as an Automated Text-to-Speech Recognition Software for Audio Description Research.* - **Zi Ye** (University of Bristol). *Exploring the audio description of humour: A multimodal corpus study in the Chinese context.*

Coffee break (11:15-11:45)

11:45–13:15	**PANEL 4. AUDIO DESCRIPTION RESEARCH** Chair: Iwona Mazur (Adam Mickiewicz University) - **Roger Johansson,** Maja Rudling, Johan Martesson, Jana Holsanova (Lund University). *Mind the Boundaries: Neurocognitive and AI-Driven Insights into Event Perception in Audio-Described Films* (15') - **Jessica Beale** (University of Westminster). *Findings from a linguistic experiment suggest that multisensory language in audio description helps listeners to generate more vivid mental imagery* (15')

	- **Harun Dalli,** Anna Jankowska, Iris Schrijver (University of Antwerp). *Audio Description Expertise and Quality: Unraveling A Conceptual Yardstick* (15') - **Inma Pedregosa** (University of Westminster). *Audio Descriptive Commentary in Football* (15') - **Alejandro Romero** (Universitat Jaume I). *The Heartbeat of Creativity in Audio Description: A Pilot Heart Rate Study on Multimodal Objectivity and Subjectivity* (10') Discussion: 30 minutes.

Lunch break (13:15-14:15) *Although lunch is not included, sandwiches will be offered to participants*

14:15–16:00	**PANEL 5. AN ACCESSIBILITY TRIP AROUND THE GLOBE** Chair: Marta Brescia (UAB) - **Blanca Arias-Badia** (Universitat Pompeu Fabra), Katixa Agirre-Miguelez (University of the Basque Country), Antía María López-Gómez (University of Santiago de Compostela), Naiara Larrakoetxea-Salgado (Basque Federation of Associations of Deaf People and Spanish Confederation of Deaf People), Marijo Deogracias-Horrillo (University of the Basque Country), Josu Amezaga-Albizu (University of the Basque Country). *Audio description on streaming platforms: A snapshot of the present language offer* (10') - **Marta Żaczkiewicz**, Anna Jankowska (University of Anwerp). *Assessing the Accessibility of Multimedia Content in Polish and EU public Institutions* (10') - **Marzieh Izadi** (University of Anwerp). *A study on the availability and reception of audio description in Iran* (15') - **Dongxue Han** (Shanghai Yier Information Co.), Yuchen Liu (Shanghai International Studies University). *The Commercialization of Audio Description in China: Present and Future Prospects through the Case Study of the Yier Company* (15') - **Suely Maciel,** Guilherme Ferreira de Oliveira, Guilherme Mori Magalhaes (São Paulo State University). *Audio Description of Spaces and Landscapes for the "SIGA – Accessible Guide of the City" App* (15') - **Federico Spoletti** (SUB-TI)**,** Annalisa Sandrelli (Università degli Studi Internazionali di Roma, SUB-TI). *INCinema: An Accessible Film Festival* (15') Discussion: 25 minutes.
16:00–18:30	**Screening of the film "Radiance/Hikari", by Naomi Kawase, with audio description.** Social event organized by SUB-TI. All ARSAD attendees are invited and registration is not needed (same room as the conference).

Informal cocktail at the venue cloister (18:30–19:30)

CONFERENCE- DAY 3: 21 MARCH 2025

08:45–09:00	Registration
09:00–10:30	**PANEL 6. BUILDING FILMIC EXPERIENCES** Chair: Sarah McDonagh (Universitat Autònoma de Barcelona) - **María José García-Vizcaíno** (Montclair State University). *Crowdsourcing in Audio Description: The Blue Description Project* (15') - **Iwona Mazur**, Julia Kaźmierczak (Adam Mickiewicz University). *Experimental Audio description as a Means of Enhancing the Cinematic Experience of Sighted Viewers* (15') - **Gert Vercauteren** (University of Antwerp). *No Sound but the AD… How does Audio Description Alter Filmic Narratives? The Case of the Flemish TV Series Nachtwacht* (15') - **Marcella Wiffler** (University of Campinas). *Creativity in audio description of the horror genre: a reception research* (15') Discussion: 30 minutes
10:30–11:00	**AD PITCHES II** (5 minutes each) Chair: Anna Matamala (UAB) - **Catarina Xavier** (Universidade de Lisboa), Cláudia Martins (Instituto Politécnico de Bragança). *Professional vs Trainee Translators and Audio Description: An Exploratory Study on Taboo Images Audio Described into Portuguese.* - **Theodora Psoma** (Aristotle Univeristy of Thessaloniki), Theodora Sazo (University of Macedonia), Christina Fragkou (University of Macedonia), Kleoniki Nerantzoglou (University of Macedonia). *What (more) can we do with words: An interdisciplinary workshop on creating Museum Audio Descriptions with the use of Generative AI.* - **Jackie Xiu Yan, Su Lin** (City University of Hong Kong). *Audio Description in the Era of AI: Challenges and Future Development* - **Gilherme Mori Magalhães**, Suely Maciel (São Paulo State University). *Characters, Scenarios and Objects Audio Description in Games: a study with Mortal Kombat 1*

Coffee break (11:00-11:30)

11:30–12:35	**PANEL 7. MUSEUM AND ART AUDIO DESCRIPTION** Chair: Blanca Arias (Universitat Pompeu Fabra) - **Martina Maggi** (University of Turin, University of Genoa). *Accessibility for All in Museums: Applying Technology to Audio Description* (15') - **María Isabel Rivas Ginel** (Dublin City University), Martin Kappus (Zurich University of Applied Sciences), Alexa Lintner (Zurich University of Applied Sciences), Joss Morkens (Dublin City University), Will Noonan (Université de Bourgogne), Lucía Pintado-Gutiérrez (Dublin City University). *A multilingual exploratory study of generative AI usage for museum audio description* (15') - **Weronika Magdalena Kostyra** (University of Warsaw). *Large Language Models as a Potential Source for Automatic Audio Description of Art in Polish. The Case of the Be My AI Tool powered by GPT-4 and GPT-4º* (15') Discussion: 20 minutes.

12:35–13:15	**PANEL 8. DANCE AD** Chair: Karen Seeley (The University of Adelaide) - **María Luján,** Marina Ramos Caro, Ana María Rojo López (Universidad de Murcia). *Hearing Dance: Poetic or Neutral AD?* (15') - **Arianna Carloni,** Sabine Braun, Dimitris Asimakoulas (University of Surrey). *Watching and Listening to Dance* (15') Discussion: 10 minutes.
13:15–13:45	**PANEL 9. DESCRIBING DIVERSITY** Chair: Marie Campbell (Red Bee Media) - **Bernd Benecke** (BR Munich). *What can I say? Describing physical aspects of people in times of inclusion and diversity* (10') - **Amanda Hiu Tung Chow,** Jackie Xiu Yan (City University of Hong Kong). *Avoiding Sexism, Racism, and Ableism in the Cantonese Audio Description of Still Human (淪落人)* (10') Discussion: 10 minutes.

Lunch break (13:45-14:45) *Although lunch is not included, sandwiches will be offered to participants*

14:45–16:00	**PANEL 10. TRAINING AND CERTIFICATION** Chair: Brett Oppegaard (University of Hawai'i at Manoa) - **Karen Seeley** (The University of Adelaide). *AD in Australia: describer accreditation* (10') - **Floriane Bardini**, Eva Espasa (Universitat de Vic). *Audio description: a transversal training tool for minoritized language education* (15') - **Caroline Martin**, Santiago Hidalgo (Montreal University). *Film students and audio description: results of a francophone inclusive media training pilot project in Québec, Canada* (15') - **Anna Jankowska,** Nina Reviers, Gert Vercauteren (University of Antwerp). *'Accessibility is Like Air - Everyone Needs it to Stay Alive' - The Impact of Accessibility Training on Awareness Levels* (10') Discussion: 25 minutes.
16:00–16:45	**PANEL 11. COGNITIVE ACCESSIBILITY AND AD** Chair: Gert Vercauteren (University of Antwerp) - **Francesca Carpenedo,** Federico Spoletti (SUB-TI). *Accessible Screenings for the Elderly: Enhancing Cognitive Stimulation and Social Well-being through the FAIR Project* (15') - **Maria Andreea Deleanu,** Constantin Orasan, Sabine Braun (University of Surrey). *Repurposing Audio Description, Audio Narrative and integrated subtitles: trialling 'Accessible Cues' with cognitively diverse audiences in the UK* (15') Discussion: 15 minutes.
16:45–17:00	**FINAL REMARKS**

WORKSHOPS

19 MARCH 9:00. ENACT pre-conference workshop on media accessibility: state of the art and future challenges

Free registration for ARSAD participants.

19 MARCH 17:30. MOSAIC project workshop. AI meets Audio Description. Exploring opportunities and risks.

The media landscape has transformed dramatically with the rise of global streaming services like Amazon Prime, Disney+, and Netflix. European broadcasters and producers recognize the need to maintain the EU market presence while reducing dependency on these platforms. Mosaic addresses this challenge by creating an AI-enhanced platform that serves as a central hub for media content creators, distributors, and consumers across Europe. In this workshop we will show the tools that have been chosen by private and public broadcasters to enhance their workflow with the aid of AI. Then we will interact with participants to understand their reactions, suggestions, and requirements.

Organiser: Pilar Orero (UAB)

All ARSAD attendees are welcomed and registration is not needed (same room as the conference).

SOCIAL EVENTS

19 March 19:00. Social Event for panel speakers. Organised by Bridge Multimedia. Event by invitation only, details to follow to panel speakers.
20 March 16:15. Screening of the film "Radiance/Hikari", by Naomi Kawase, with audio description. Organised by SUB-TI. All ARSAD attendees are invited and registration is not needed (same room as the conference).
20 March 18:30. Informal cocktail dinner All ARSAD attendees are invited and registration is not needed (at the venue cloister).

KEYNOTE SPEAKER: SABINE BRAUN

Bridging the Accessibility Gap? The role of AI in the future of audio description generation

Access to digital media is essential for fostering an inclusive society and reducing inequalities, in line with United Nations Sustainable Development Goals 10 and 16. However, as the demand for accessible media content rises—driven by the continuous increase in audiovisual content and evolving accessibility legislation—the traditional workflows for creating audio description (AD) are increasingly strained. Moreover, while human-generated AD remains the gold standard for quality, its delivery is more challenging in real-time or near-real-time settings. To address these challenges, there has been growing interest in leveraging computer vision, machine learning, and generative artificial intelligence (AI) methods to assist or automate AD creation (Braun & Starr, 2019, 2022).

Drawing on advances in automated video scene understanding, video captioning and speech technologies, and building on the instruction-following capabilities of large language and multimodal models, the computer science and engineering research communities have made progress not only in automating character identification and AD script generation but also in automatic gap detection for AD insertion and the voicing of AD scripts.

In video scene understanding, expanding image-text representations across multiple frames has enabled character tracking and the recognition of actions and interactions (Chu et al., 2024; Han et al., 2023a, 2023b; Ye et al., 2024), improving the continuity of descriptions compared to previous approaches (Braun et al., 2020). For AD generation, two-step approaches involving dense video captioning and summarisation have produced richer and more detailed descriptions than previous techniques (Chu et al., Shen et al., 2023; Wang et al., 2021), while complex multimodal pipelines have integrated subtitles, scripts, audio tracks, and contextual data in the AD generation task for greater accuracy and relevance of the descriptions (Campos et al., 2020; Chu et al., 2024; Han et al., 2023a, 2023b; Wang et al., 2024; Ye et al., 2024).

However, much of this research remains at an experimental stage. User engagement is limited, and generative AI poses risks related to accuracy and bias (Bergin & Oppegaard, 2024), raising important questions about its viability for AD. The limitations of AI regarding contextual awareness and linguistic nuance can result in misleading or inappropriate outputs, undermining meaningful accessibility and

user trust. Therefore, the role of AI in future AD workflows requires further scrutiny.

This presentation will first provide an overview of the current AD automation landscape and initial findings from an evaluation of recent AI models, assessing their potential and limitations. It then examines broader implications for the role of AI in AD workflows, reflecting on alternatives to full automation, from AI-assisted AD generation with post-editing (Braun et al., 2021; Chang et al., 2024) to collaborative AD production through upskilling of volunteers (Natalie et al., 2021) and integrative AD (Fryer, 2018). Finally, it briefly considers whether incorporating human knowledge about AD into AI development (Johansson et al., 2024; Starr et al., 2020) may enhance the safety and reliability of AI-generated AD.

Overall, this presentation aims to highlight the benefits of human-centric approaches to AI in AD generation (Capel & Brereton, 2023), focussing on how AI and human expertise can complement each other in delivering accessible content at scale without compromising quality or ethical standards.

References

Bergin, D., Oppegaard, B., 2024. Automating Media Accessibility: An Approach for Analyzing Audio Description Across Generative Artificial Intelligence Algorithms. Technical Communication Quarterly 1–16. https://doi.org/10.1080/10572252.2024.2372771

Braun, S., Starr, K., 2019. Finding the Right Words: Investigating Machine-Generated Video Description Quality Using a Corpus-Based Approach. Journal of Audiovisual Translation 2, 11–35. https://doi.org/10.47476/jat.v2i2.103

Braun, S., Starr, K., 2022. Automating Audio Description, in: Taylor, C., Perego, E. (Eds.), The Routledge Handbook of Audio Description. Routledge, London, pp. 391–406. https://doi.org/10.4324/9781003003052-30

Braun, S., Starr, K., Delfani, J., Tiittula, L., Laaksonen, J., Braeckman, K., Van Rijsselbergen, D., Lagrillière, S., Saarikoski, L., 2021. When Worlds Collide: AI-Created, Human-Mediated Video Description Services and the User Experience, in: Lecture Notes in Computer Science. Springer, Washington D.C./Virtual, pp. 147–167. https://doi.org/10.1007/978-3-030-90328-2_10

Braun, S., Starr, K., Laaksonen, J., 2020. Comparing human and automated approaches to visual storytelling, in: Braun, S., Starr, K. (Eds.), Innovation in Audio Description Research. Routledge, London,

pp. 159–196. https://doi.org/10.4324/9781003052968-9

Campos, V.P., de Araújo, T.M.U., de Souza Filho, G.L., Gonçalves, L.M.G., 2020. CineAD: a System for Automated Audio Description Script Generation for the Visually Impaired. Universal Access in the Information Society 19, 99–111. https://doi.org/10.1007/s10209-018-0634-4

Capel, T., Brereton, M., 2023. What is Human-Centered about Human-Centered AI? A Map of the Research Landscape, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. ACM, Hamburg Germany, pp. 1–23. https://doi.org/10.1145/3544548.3580959

Chang, R.-C.C., Sheshadri, S., Guo, A., Hara, K.-T.K., Natalie, R., Chang, R.-C.C., Sheshadri, S., Guo, A., Hara, K.-T.K., 2024. Audio Description Customization. ASSETS 2024 - Proceedings of the 26th International ACM SIGACCESS 1, 19. https://doi.org/10.1145/3663548.3675617

Chu, P., Wang, J., Abrantes, A., 2024. LLM-AD: Large Language Model based Audio Description System. https://arxiv.org/abs/2405.00983v1

Fryer, L., 2018. Staging the Audio Describer: An Exploration of Integrated Audio Description. Disability Studies Quarterly 38. https://doi.org/10.18061/dsq.v38i3.6490

Han, T., Bain, M., Nagrani, A., Varol, G., Xie, W., Zisserman, A., 2023a. AutoAD: Movie Description in Context. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2023-June, 18930–18940. https://doi.org/10.1109/CVPR52729.2023.01815

Han, T., Bain, M., Nagrani, A., Varol, G., Xie, W., Zisserman, A., 2023b. AutoAD II: The Sequel – Who, When, and What in Movie Audio Description. Proceedings of the IEEE International Conference on Computer Vision 13599–13609. https://doi.org/10.1109/ICCV51070.2023.01255

Johansson, R., Rastegar, T., Lyberg-Åhlander, V., Holsanova, J., 2024. Event Boundary Perception in Audio Described Films by People Without Sight. Applied Cognitive Psychology 38, e4228. https://doi.org/10.1002/acp.4228

Natalie, R., Loh, J., Tan, H.S., Tseng, J., Chan, I.L.Y.-R., Jarjue, E.H., Kacorri, H., Hara, K., 2021. The Efficacy of Collaborative Authoring of Video Scene Descriptions, in: Proceedings of the 23rd International ACM SIGACCESS. ACM, New York, NY, USA, pp. 1–15.

https://doi.org/10.1145/3441852.3471201

Pavel, A., Reyes, G., Bigham, J.P., 2020. Rescribe, in: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. ACM, New York, NY, USA, pp. 747–759. https://doi.org/10.1145/3379337.3415864

Shen, X., Li, D., Zhou, J., Qin, Z., He, B., Han, X., Li, A., Dai, Y., Kong, L., Wang, M., Qiao, Y., Zhong, Y., 2023. Fine-grained Audible Video Description, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 10585–10596. https://doi.org/10.1109/CVPR52729.2023.01020

Starr, K.L., Braun, S., Delfani, J., 2020. Taking a Cue from the Human: Linguistic and Visual Prompts for the Automatic Sequencing of Multimodal Narrative. Journal of Audiovisual Translation 3, 140–169. https://doi.org/10.47476/jat.v3i2.2020.138

Wang, H., Tong, Z., Zheng, K., Shen, Y., Wang, L., 2024. Contextual AD Narration with Interleaved Multimodal Sequence. https://arxiv.org/abs/2403.12922v1

Wang, Y., Liang, W., Huang, H., Zhang, Y., Li, D., Yu, L.-F., 2021. Toward Automatic Audio Description Generation for Accessible Videos, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 1–12. https://doi.org/10.1145/3411764.3445347

Ye, X., Chen, J., Li, X., Xin, H., Li, C., Zhou, S., Bu, J., 2024. MMAD: Multi-modal Movie Audio Description, in: Calzolari, N., Kan, M.-Y., Hoste, V., Lenci, A., Sakti, S., Xue, N. (Eds.), Proceedings of the 2024 LREC-COLING 2024), Torino, Italia, pp. 11415–11428.

Dr Sabine Braun is full Professor of Translation Studies, Director of the Centre for Translation Studies at the University of Surrey, and a Co-Director of the Surrey Institute for People-Centred AI. Her research explores human-machine interaction and integration in different forms cross-lingual and cross-modal mediation (e.g., interpreting, audio description) to improve access to critical information, media content, and essential public services. In relation to media access, she has a long-standing interest in exploring the cognitive, discursive and narratological foundations of audio description. As a Workpackage leader in the European Horizon 2020 Project MeMAD (Methods for Managing Audiovisual Data, 2018-21), she collaborated with computer vision experts to investigate the feasibility of (semi-)automating the creation of AD to improve access to media content. She co-edited 'Innovations in Audio Description Research' (with Kim Starr; Routledge, 2021). In 2024, she launched the Leverhulme Trust-funded Doctoral Training Network on AI-

Enabled Digital Accessibility (ADA), which includes AI-enabled audio description as one of its main research strands. She is supervising various PhD projects on audio description, is involved in a proof-of-concept project on incorporating audio description into video games and is regularly advising industry and third-sector organisations on the use of AI in audio description generation.

PANEL 1. TECHNOLOGY: INDUSTRY PERSPECTIVES

AI and AD: For Good or Ill?

Joel Snyder
jsnyder@audiodescribe.com
Audio Description Associates, LLC; Audio Description Project of the American Council of the Blind

When I teach audio description (AD) at sessions around the world, I focus a great deal on the crafting of the language used--and most AD is written to be heard. Trained description writers and voice talents trained in voicing techniques for AD are critical to the success of the effort to translate a visual image to the spoken word.

It seems as though we are on the cusp of AI dominance in all manner of endeavor. Speech synthesis is already employed by some companies who produce AD for broadcast television. But the writing of audio description is dependent on an understanding and thorough analysis of the work to be described. Similarly, the appropriate voicing of AD is done with nuance, attention to the images on screen, and an understanding of the phrasing used in the writing of the description. Indeed, in 2021, the American Council of the Blind passed a resolution noting its "full support for … the use of human voices in the voicing of audio description for cinema and narrative video or streaming." https://www.acb.org/2021-resolutions#Resolution2122

Can the spread of AI/speech synthesis be stopped or forestalled? Should it be?

The presentation will offer AI inquiries and responses posed to ChatGPT regarding the effectiveness of AI's drafting and voicing of AD as well as a comparison of human and AI-generated AD excerpts.

AI apps are inevitable and will surely bring great advances to humanity. But—at least for the foreseeable future—experienced and *human* writers and voice talents are key to effective AD.

Dr. Joel Snyder is known internationally as one of the world's first "audio describers," a pioneer in the field of Audio Description. Since 1981, Dr. Snyder has introduced audio description techniques in over 40 states in the U.S. and 64 countries; he has made thousands of live events, media projects and museums accessible. Dr. Snyder was named a Fulbright Scholar to train audio describers in Greece and Ukraine over a four-week period and recently trained describers in Bulgaria and Uzbekistan. The American Council of the Blind published

Dr. Snyder's book, The Visual Made Verbal – A Comprehensive Training Manual and Guide to the History and Applications of Audio Description, available as an audio book and in Braille from the Library of Congress and in English, Polish, Russian, Portuguese, Spanish, Italian, and Chinese print editions.

Additive Prompting: A Tool for AI Assisted Audio Description

Matthew Kaplowitz, Wendy Sapp
mkaplowitz@bridgemultimedia.com
Bridge Multimedia

With the rapidly expanding pace of Generative Artificial Intelligence as a tool for accessibility, Additive Prompting is a rich area for innovation to enhance the experiences of learners, provide stimulation, build curiosity, and inspire creativity. Additive prompting is a method for creating organized and well-structured prompts for artificial intelligence (AI) to follow. By integrating well-established Additive Prompting strategies within AI into accessibility practices, children and adults of all abilities can create and enjoy accessible products.

Additive Prompting begins by conceptualizing approximately five prompt categories needed to create an item. Starting with a simple idea, the user slowly builds up the complexity of their creation. If a single complex prompt is input, AI struggles to include all the features or prioritize the aspects of the prompt. By dividing the instructions given to the AI engine into categories, the AI is more easily able to follow the instructions and the user can easily identify and correct any omissions or errors. For example, to create an AI generated image, common categories are (1) subject, (2) style, (3) lighting or colors, (4) background, and (5) extras. A user could create an image by inputting characteristics in each category such as in this example: (1) dog, (2) illustration, (3) neon colors, (4) in a field of wildflowers, and (5) chasing butterflies. Hopefully, the resulting AI generated image would closely match the users intended goal.

While Additive Prompting is frequently used for AI generated images, the principle of creating organized and well-structured prompts can be used for any AI generated product, including but not limited to poems, informational essays, 3-D printed items, graphs or maps, and videos. When structured properly and integrated into accessible and assistive technology, anyone can use Additive Prompting to produce a creative work (such as the previous example of a drawing of a dog) or demonstrate their understanding of an academic concept (such as creating a video to illustrate the water cycle and its impact on climate).

Audio Description, Accessibility, and Generative AI: Images, videos, and other graphic products created through using Additive Prompting require image description or audio description to be accessible. By using the original prompts, AI has sufficient data and metadata to produce meaningful and accurate descriptions with minimal human intervention. To meet the needs of the intended audience, the creator

can then provide additional prompting categories as appropriate such as comprehension level or length of the description, and the corresponding image description or audio description can be generated. Human checking and revisions to the AI generated description are always recommended, but beginning with an AI generated description can facilitate the creation of description. With Additive Prompting, anyone can create an image or video and, with only minor additional effort, create an image description or audio description. Integrating Additive Prompting with accessibility requirements opens new horizons for audio description and lowers the hurdles new creators face when attempting to include audio description in their creations.

Matt Kaplowitz is founder, chief executive officer, and chief creative officer of Bridge Multimedia, an accessibility enterprise based in New York. Since 2000, Mr. Kaplowitz has led a research and development team, implementing universally accessible technology for television and electronic media and creating a technology and workflow protocol for handling and tracking large volumes of multimedia data. Mr. Kaplowitz has been involved in major literacy, science, math, and English Language Learner curriculum series for Scholastic, Oxford University Press, Houghton Mifflin and Pearson Education. He has been at the forefront of the development of surround sound for television and for digital cinema in association with Dolby Laboratories and Boeing Satellite Services. Mr. Kaplowitz's work has earned him several awards, including an Emmy, a Grammy, a Peabody, a Clio, a Cable Ace Award, and an award from Action for Children's Television.

Wendy Sapp, Ph.D., is Chief Research Officer and Senior Project Director at Bridge Multimedia, an accessibility enterprise based in New York City. Since 2006, Dr. Sapp has led audio description teams that have created thousands of hours of audio description for educational television programming for children and adolescents. During this time, she also led additional accessible technology teams, including those that created five award winning accessible digital games, multiple accessible virtual conferences, and several research studies in media accessibility. Dr. Sapp serves on the Subject Matter Expert Committee for Audio Description of the Academy for the Certification of Vision Rehabilitation and Education Professionals, which is developing certification standards for Audio Description Specialists.

Using AI for audio description generation

Eveline Ferwerda
eveline@scribit.pro
Scribit.Pro

Research into the (im)possibilities of AI for the production of audio description for drama content

Scribit.Pro recently conducted a study for the Bartiméus Fund in which we mapped the (im)possibility of artificial intelligence for the production of audio description for video content. Those results are promising. Especially when it comes to short fragments, various AI providers are able to formulate an accurate representation of what is shown on the screen. But what is possible when it comes to drama content, in which a story unfolds over a longer period of time? Can AI create characters and connections between scenes?

AI developments are moving at lightning speed, and this summer a new version of Chat GPT will be released that focuses entirely on moving images (creating/recognizing). Scribit.Pro would like to conduct follow-up research into the (im)possibilities of AI for the production of audio description for film material to determine whether efficiency gains can be achieved in this regard.

The volume of audio description production worldwide is estimated at 360 million euros in 2030[1]. It is expected that this market will grow further if this becomes a legal obligation for many content providers. It is therefore not surprising that both producers and customers are looking for ways to make the production process more efficient and cheaper. The possibilities and application of AI seem to be the obvious answer. But can an algorithm actually convey a story smoothly and clearly to the viewer, despite the limited time available for this? There is only space between the spoken dialogues. Can AI determine what is most essential and can it provide the viewer with support in interpreting the entire narrative?

Our initial research shows that AI achieves better results if more information and/or context feeds into your 'prompt'; the task that AI must carry out. What happens if, in addition to the video material, we also add the film script to the request? And how do these results compare to a human audio describer of the same material?

---

[1] Source: Audio Description Services Market To USD 0.39 Billion By 2030, Analysis By Exactitude Consultancy https://menafn.com/1108052965/Audio-Description-Services-Market-To-USD-039-Billion-By-2030-Analysis-By-Exactitude-Consultancy)

The research consists of describing 5 short films by both a professional audio describer and the various AI models. What results does that produce? What experiments can we perform on this in 'prompt engineering'? In short, is AI able to contribute to the production process of audio description in the future?

Research questions:

- How well can we take advantage of developments in AI in audio description production?
- What do the new models in AI bring to the production of audio description for drama content?
- How well does the improved prompt engineering work – providing subtitle/script information

The research will start September 1st, 2024, and will be done in January 2025. It is financed by Bartimeus Fund, The Fund for Film producers and the Dutch broadcasters innovation fund.

Eveline Ferwerda came in contact with Accessibility working as a project manager for Bartiméus Foundation, a Netherlands-based NGO. With her background in the cinema industry, Eveline strives to incorporate people with various visual limitations into the cinematic experience to the fullest extent possible. She created the scribit.tv platform in 2018 and found funds to evolve Scribit into a successful social enterprise. In 2020 the Scribit.Pro playform was launched, a one-stop-shop online video accessibility tool.

PANEL 2. TECHNOLOGY: RESEARCH PERSPECTIVES

A Strange New World: Behind the Scenes of Designing Novel Gen-AI Approaches to Audio Description

Brett Oppegaard
brett.oppegaard@hawaii.edu
University of Hawaii at Manoa

Tech companies have been for decades pumping out overwhelming quantities of inaccessible digital media while others have been following in the wake, promising silver-bullet accessibility solutions to remedy such widespread access issues. Until November 2022, when OpenAI released the GenerativeAI (GenAI) ChatGPT system to the public, the remedies had been mostly vaporous.

But ChatGPT and its GenAI brethren in just over a year have suddenly changed the media environment, including upending common approaches to accessibility. In recent months, fueled by the release of game-changing upgrades, such as ChatGPT-4o, emerging GenAI-generated Audio Description (GenAI-AD) has brought back to life promises of automated and widespread advances in accessibility for people who cannot see or see well. In response, our research team set out to thoroughly test these claims and have indeed documented rapid advancements and quickly emerging opportunities for Audio Description innovation.

This paper and presentation will report about some of those GenAI-AD experiments and tests of these systems from their earliest public forms to today, including a novel double-rendering method that asks GenAI-AD engines to describe a simple portrait of a person and then returns these generated texts into GenAI-AD engines for visualizations of what they earlier had described, revealing ongoing insights about GenAI efficacies, ethics, and biases. This method provides a straightforward and simple test of the systems, which anyone can do, forcing the GenAI-AD to prove its abilities through the creation of a common Audio Description product, which then can be evaluated for qualitative aspects and compared to previous iterations. Another part of this presentation will report about the creation of GenAI bots designed to coach audio describers through the early steps of creating complicated descriptions of static media, ranging from photographs to maps to tables and charts. These AI Guidedogs, as we called them, each have distinct personalities and rhetorical coaching strategies, and they are not intended to replace people but to help describers become better at what they do and to consistently perform better than their AI-AD counterparts. That strategic approach we took aims to preserve the place of people in the describing

process. But that's because, to put it bluntly, if GenAI continues to develop and expand at the current exhilarating pace, it will radically disrupt and possibly create an existential threat to many of the legacy Audio Description systems currently in operation. To keep up, human describers will need to get better, too.

These technical developments simultaneously open many new opportunities as well, particularly for those in the right position and with the right mindset to adapt and to realize these potentials. These GenAI-AD systems already can work exponentially faster than human-pace, without needs for breaks or vacations, without management issues, or writer's block. Humans already cannot compete with the quantity of information that AIs can produce, but even more unsettling for audio describers, the human-led projects we employ today might also eventually become less efficient and effective in a qualitative sense as well. So, what are people going to do about that? We discuss.

Brett Oppegaard, Ph.D., researches media-production processes and products at intersections of Technical Communication, Rhetoric, Disability Studies, Digital Inequalities, and Journalism. He primarily studies those interests within mobile contexts, including creating novel research tools through interactive mobile media, mobile apps, and mobile technologies.

Emotions offstage: diversity and violence described by AI

Marina Ramos Caro, Purificación Meseguer Cutillas
marinaramos@um.es
University of Murcia, Spain

In the past decade, there has been a timid shift in audio description (AD) research towards process evaluation. The wide variability in AD outputs from the same AD task has been one of the main findings, observed in several studies in the form of a corpus of AD scripts created by professionals and students in English, Catalan, and Spanish (Matamala, 2019), to test the influence of creativity (Ramos & Rojo, 2020) and ideology (Ramos & Meseguer, 2023; Meseguer & Ramos, 2023) on the AD output, and most recently, to analyze the rendering of generative artificial intelligence (GenAI) tools (Bergin & Oppegaard, 2024).

To further analyze this phenomenon, we decided to delve deeper into the challenges of describing diversity by replicating our previous study (Ramos & Meseguer, 2023) with GenAI AD tools. We analyzed the AD output resulting from the description of the same 10 images from our previous study (Ramos & Meseguer, 2023) depicting different types of diversity (sexual orientation, gender, functional diversity, racial and ethnic diversity) and sexual violence. However, this time the output was created by 7 different GenAI tools: Microsoft Azure Cognitive Services, Google Cloud Vision API, Amazon Rekognition, IBM Watson Visual Recognition, Clarifai, Microsoft's CaptionBot, and DeepAI Image Captioning API. Our results confirm the biases already observed by Bergin & Oppegaard (2024) and pose new research and ethical questions regarding the description of diversity.

References

Bergin, D., & Oppegaard, B. (2024). Media accessibility: An approach for analyzing audio description across generative artificial intelligence algorithms. *Technical Communication Quarterly*, *33*(1), 1–16. https://doi.org/10.1080/10572252.2024.2372771

Matamala, A. (2019). The VIW corpus: Multimodal corpus linguistics for audio description analysis. *RESLA. Revista Española de Lingüística Aplicada*, *32*(2), 515–542.

Meseguer Cutillas, M., & Ramos Caro, M. (2023, July 5–7). M4All. Paper presented at the M4All International Conference, University of Antwerp, Antwerp, Belgium.

Ramos Caro, M., & Meseguer Cutillas, P. (2023, April 20–21). The role of ideology on the audio description process. Paper presented at the Advanced Research Seminar on Audio Description, Universitat Autònoma de Barcelona, Barcelona, Spain.

Ramos Caro, M., & Rojo López, A. M. (2020). Analysing the AD process: Creativity, accuracy and experience. *Journal of Specialised Translation*, *34*, 212–232.

Marina Ramos Caro is Lecturer in Translation at the University of Murcia (Spain). She is particularly interested in the influence of emotions and personality in the process and reception of Audio Description and translation. Dr. Ramos has been awarded several international grants and presented her work in international conferences. Her work has been published in specialized journals such as Journal of Pragmatics, MonTi, Translation Spaces, The Translator, Perspectives, JosTrans and Frontiers. As a professional translator, she has worked within the fields AVT and transcreation. She is the founder of Latrium, the Laboratory for Inclusive Translation of the University of Murcia (IG: Latri_um).

Purificación Meseguer Cutillas is a Lecturer at the Department of Translation and Interpreting at the University of Murcia. Her main research interests are the relationship between translation and censorship, and the impact of emotions and personality factors in translation. She has written a book (Sobre la traducción de libros al servicio del franquismo. Peter Lang, 2015) and published many articles in specialised national and international journals. As a professional translator she has translated a wide range of books for publishing houses such as Tusquets, RBA or Random House, an experience which helps her to bring the professional world into the classroom.

An exploratory analysis of LLM-based machine-translated audio description scripts in the French-German language pair

[1]Nicole Lüthi, [1]Alexa Lintner, [2]Lukas Fischer, [1]Martin Kappus, [2]Sarah Ebling
luethinicole@hispeed.ch
[1]Zurich University of Applied Sciences
[2]University of Zurich

(Semi)automated methods in audio description (AD) production, e.g. Large Language Model (LLM)-based machine translation (MT) of AD scripts can address the rising demand for ADs due to reinforced legal requirements (cf. Braun & Starr 2022). This is particularly true for multilingual countries such as Switzerland where AD scripts are currently being created ex novo in three national languages. Previous studies indicate that the use of MT could be an alternative production method in AD and recommend further investigation into post-editing effort (Fernández-Torné 2016; Fernández-Torné & Matamala 2016; Matamala & Ortiz-Boiz 2016; Vercauteren et al. 2021). Considering recent developments in artificial intelligence (AI) and the use of LLM-based MT, this becomes all the more important.

In line with common practice in (human) MT quality assessment (cf. e.g. Bentivogli et al. 2018), we conducted a source-based direct assessment with professional audio describers to evaluate machine-translated AD segments in the French-German language pair (Lüthi 2024). AD segments had been translated with OpenAI's GPT-4 model. AD professionals evaluated three different translation versions: 1) generated with textual input only; 2) generated with both textual and visual input, 3) the latter post-edited by ChatGPT-4. Evaluators were asked to assign adequacy, fluency and usefulness scores to every AD segment without knowing how each translation version had been produced. Additionally, we applied a qualitative approach and analyzed selected AD segments using the harmonized DQF-MQM error typology (cf. Lommel 2018). Our results suggest that LLM-based, human-assisted MT can be a viable alternative method to produce ready-for-voicing AD scripts.

References

Bentivogli, L., Cettolo, M., Marcello, F., & Federmann, C. (2018). Machine translation human evaluation: An investigation of evaluation based on post-editing and its relation with direct assessment. In *Proceedings of the 15th International Workshop on Spoken Language Translation (IWSLT 2018)* (pp. 62–69).

Braun, S., & Starr, K. (2022). Automating audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio*

description (pp. 391–406). Routledge. https://doi.org/10.4324/9781003003052-30

Fernández-Torné, A. (2016). Machine translation evaluation through post-editing measures in audio description. *inTRAlinea, 18*. https://www.intralinea.org/archive/article/2200

Fernández-Torné, A., & Matamala, A. (2016). Machine translation in audio description? Comparing creation, translation and post-editing efforts. *SKASE Journal of Translation and Interpretation, 9*(1), 64–87.

Lommel, A. (2018). Metrics for translation quality assessment: A case for standardising error typologies. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 109–127). Springer. https://doi.org/10.1007/978-3-319-91241-7_6

Matamala, A., & Ortiz-Boix, C. (2016). Accessibility and multilingualism: An exploratory study on the machine translation of audio descriptions. *TRANS, 20*, 11–24. https://doi.org/10.24310/TRANS.2016.v0i20.2059

Lüthi, N. (2024). KI-gestützte (maschinelle) Übersetzung von Audiodeskription. Eine quantitative und qualitative Untersuchung der Sendung *Passe-moi les jumelles* [Unpublished master's thesis]. ZHAW Zurich University of Applied Sciences.

Vercauteren, G., Reviers, N., & Steyaert, K. (2021). Evaluating the effectiveness of machine translation of audio description: The results of two pilot studies in the English-Dutch language pair. *Revista Tradumàtica – Tecnologies de la Traducció, 19*, 226–252. https://doi.org/10.5565/rev/tradumatica.288

Nicole Lüthi has completed her MA degree in Applied Linguistics (Specialisation in Professional Translation) at the ZHAW in summer 2024. Her MA thesis examines AI-assisted (machine) translation of audio description scripts (AD scripts) from French to German. In 2022, she also obtained her BA degree in Applied Languages at the ZHAW. For her undergraduate thesis she conducted a comparative analysis of English, German and Spanish AD scripts of the Netflix original series Bridgerton.

Alexa Lintner is a Research Associate in Accessibility Studies at the Zurich University of Applied Sciences, where she completed her MA in Applied Linguistics in 2018. Her Master's thesis was awarded the 2018 Lionbridge Prize. Alexa's research interests focus on audio description and text simplification. She is currently working on the large-scale, Innosuisse-funded research project entitled "Inclusive

Information and Communication Technologies" (IICT) aiming to radically change production workflows in accessible communication, i.e. from human-based to AI-assisted to fully automated. Alexa co-founded the Zurich-based traduko cooperative, where she works as a part-time translator, audio describer and subtitler.

Lukas Fischer is an Academic Associate working at the Language, Technology and Accessibility chair at the University of Zurich. His current work focuses on audio description translation and text simplification within the "Inclusive Information and Communication Technologies" (IICT) project. Previously, Lukas contributed to the Bullinger Project, where he specialized in machine translation for Latin.

Martin Kappus has been a lecturer at the ZHAW School of Translation and Interpreting for 15 years. He holds a PhD from Stony Brook University. Before joining ZHAW, he gained valuable experience in the translation industry as a consultant for a CAT-Tool manufacturer and a large language service provider. His research and teaching interests focus on language technology, with particular emphasis on translation technology and accessibility studies. He also provides live audio descriptions for soccer matches in Switzerland.

Sarah Ebling is Full Professor ad personam for Language, Technology and Accessibility at the University of Zurich. Her research focuses on language-based assistive technologies in the context of persons with disabilities. The focus is on basic and application-oriented research into the development and reception of these technologies. Her research takes place in the context of hearing and visual impairments, cognitive impairments, and language disorders, dealing with sign language technologies, automatic text simplification, technologies for the audio description process and computer-aided language sample analysis. Transversal to this is a focus on multimodality, e.g., with different production modalities in sign languages (manual and non-manual components) or with text and video as part of automatic translation of audio descriptions.
Sarah is involved in international and national projects and is PI of a large-scale Swiss innovation project entitled "Inclusive Information and Communication Technologies" (2022-2026; https://www.iict.uzh.ch/).

Exploring the Impact of AI Voices in Erotic Film Audio Description: A Gender-Based Study

Blanca Serrano Alcaraz, Ana María Rojo López, Marina Ramos Caro
blanca.serrano@um.es
Universidad de Murcia, Spain

While AD is widely utilized in mainstream media (López Rubio, 2024; Marra, 2023; Jankowska et al., 2017), its application within adult entertainment remains under-researched, with a few exceptions (e.g., Ramos et al., 2021; Rojo et al., 2021). This research aimed to understand whether the gender of the narrator of the audio description impacts user experience and accessibility in explicit material.

This study investigates the impact of artificial intelligence (AI)-generated voices in audio description, focusing on the effect of narrator gender on audience reception. Two AI-generated voices, one female and one male, were selected from a voice bank for their superior expressiveness achieved through advanced AI techniques and were used to narrate audio descriptions for explicit content. The audience, consisting primarily of heterosexual men and women, was evaluated to determine how the gender of the AI narrator influenced their affective reactions and sexual arousal. Additionally, a question regarding the quality of the narration was included, with a view to a potential future study which analyzes the difference in emotional reception of AI vs. human-narrated AD.

Reception measures included both self-report and physiological measures. Self-report data were collected with three questionnaires: The Tactile Version of the Self-Assessment Manikin (Iturregui, G., Méndez, J., 2019), The Sexual Inhibition / Sexual Excitation Scale, developed by Carpenter et al. (2010), and the Ratings of Sexual Arousal (Mosher, 2011). Physiological data on participants' galvanic skin response and heart rate was collected with the Shimmer3 GSR+ device. Salivary cortisol was also collected with the Salivette® Cortisol.

Using a between-subject comparison design, a total sample of 30 visually impaired participants was distributed into two different groups, one for each version of AD. The presentation of the stimuli was randomized and balanced across participants. Results revealed interesting differences in participants' levels of affective reactions and sexual arousal between the two narrations of AD. The multimethod approach adopted provided a comprehensive understanding of how the narrator gender of AD influenced users' experience and perception. Our findings contribute to the ongoing dialogue about the

development of AI in accessible media, offering valuable insights into effective practices for AD.

References

López Rubio, M. (2024). Descripción y recepción de la audiodescripción de comedias en Netflix España para personas con discapacidad visual: Una propuesta de mejora [Unpublished doctoral dissertation]. Universitat de València.

Ramos Caro, M., Espín López, L., & Rojo López, A. (2021). The psychophysiological impact of audio described porn. *Onomázein, 8*, 104–127.

Rojo López, A. M., Ramos Caro, M., & Espín López, L. (2021). Audio described vs. audiovisual porn: Cortisol, heart rate and engagement in visually impaired vs. sighted participants. *Frontiers in Psychology, 12*, Article 661452. https://doi.org/10.3389/fpsyg.2021.661452

Marra, L. (2023). The expression of emotions in the Spanish and Italian filmic audio descriptions of *The King's Speech*. *Journal of Audiovisual Translation, 6*(2), 33–54. https://doi.org/10.47476/jat.v6i2.2023.247

Jankowska, A., Ziółko, B., Igras-Cybulska, M., & Psiuk, A. (2017). Reading rate in filmic audio description. *Rivista internazionale di tecnica della traduzione, 19*, 75–97.

Blanca Serrano Alcaraz graduated with a degree in Translation and Interpreting (English) from the Universidad de Murcia in 2017, which included an Erasmus+ exchange at the University of Essex, in the United Kingdom. She later finished a Master's degree in Audiovisual Translation at Universidad Autónoma de Barcelona. During her academic career, she gained practical experience as a translator and reviewer for marketing, localization, audiovisual and accessibility companies. She also specialized in professional editing and software localization. Professionally, she has accumulated four years of experience as a translation project manager and as an account manager in the translation industry. For the last three years, she has been working as a freelance translator and is pursuing a PhD in the field of media accessibility. In 2023, she began teaching as an adjunct professor in the Translation and Interpreting degree at Universidad de Murcia.

Ana María Rojo López is a full professor of Translation and Interpretation at UMU. She coordinates the research group "Translation, Didactics, and Cognition" and is a member of the Thematic Network on Empirical and Experimental Research and the

International Laboratory for Multilectal Mediated Cognition and Communication. Her research focuses on the areas of translation and cognition, with a central emphasis on research methodology and the role of emotions and personality factors in the translation process. She has published numerous books and scholarly works in prestigious national and international publishing houses and journals.

Marina Ramos Caro is associate professor of Translation and Interpreting at the University of Murcia (Spain). As a researcher, she is particularly interested in the influence of emotions and personality on the process and reception of audio description and translation. She has received several international grants and has presented her work in numerous prestigious journals, as well as at conferences around the world. As a professional translator, she has worked mainly in the fields of subtitling and transcreation from English, German and Portuguese into Spanish. She is the founder of Latrium, the Inclusive Translation Laboratory of the University of Murcia, a project dedicated to accessibility in the performing arts for people with sensory disabilities.

AIAD Project: Making Video Games Accessible for Players with Visual Impairments Using Dynamic AI Audio Description

[1]Xiaochun Zhang, [2]Ben Ackland, [2]Thomas Keane
xiaochun.zhang@ucl.ac.uk
[1]University College London
[2]Meaning Machine

Video gaming has become a globally popular form of entertainment, yet visually impaired individuals often face challenges accessing most games. Audio Description (AD) is a proven method for enhancing content accessibility for people with visual impairments. The current challenge is to develop technology that can replicate human-provided AD at scale. This presentation will discuss the AIAD project (2024-2025), funded by Higher Education Innovation, which explores the potential of AI-enabled audio description to make video games accessible on a large scale. The project investigates the integration of Large Language Models (LLMs) with live game data to provide 'Dynamic AI AD'. This innovative approach aims to determine whether Dynamic AI AD can enhance both the playability and enjoyment of games for visually impaired players. The research examines the effectiveness of this technology in delivering real-time, contextually accurate AD that adapts to the dynamic nature of video games. Additionally, the project assesses whether Dynamic AI AD offers a scalable and cost-efficient solution for widespread adoption in the gaming industry. By leveraging AI, the project seeks to create a more inclusive gaming environment, ensuring that visually impaired players can fully engage with and enjoy video games.

This presentation will feature a demonstration of the Dynamic AI Audio Description (AI AD) integrated into a video game. The demo will showcase the AI's ability to interpret live game data and deliver descriptions that keep pace with the dynamic nature of the game environment. Following the demo, the presentation will delve into the research findings from the AIAD project. Key insights will be shared on the effectiveness of Dynamic AI AD in improving both the playability and enjoyment of video games for visually impaired users. The discussion will cover how the technology was tested, the metrics used to evaluate its performance, and the feedback received from participants. Additionally, the presentation will explore the broader implications of this research. This includes the potential for Dynamic AI AD to be scaled and adopted widely within the gaming industry, offering a cost-efficient solution for making games more accessible. The implications for future developments in AI-driven accessibility tools will also be considered, highlighting how this project could pave the way for more inclusive gaming experiences.

Dr Xiaochun Zhang is a Lecturer in Translation Studies at the University College London (UCL), United Kingdom. Her research interests lie primarily in video game localisation and accessibility, fan audiovisual translation, and language technology. She is the principal investigator of the AD4Games and TransAD4Games projects, which investigate the application of audio description in video games to improve accessibility.

Meaning Machine is start-up company, building on Ben Ackland's 20 years' experience building innovative software products and businesses, and Thomas Keane's work on natural language gameplay systems, including the world's first-ever "First Person Talker" game.

PANEL 3. AD PROCESSES: LOOKING INTO THE FUTURE

Modular Audio Description: a novel approach to AD creation and delivery

Olga Davis, Dimitris Asimakoulas, Sabine Braun
o.davis@surrey.ac.uk
University of Surrey

In the field of Media Accessibility, audio description (AD) is the traditional method of making visual content accessible in a verbal format for visually impaired users. Traditional screen AD is highly time constrained, as the commentary is delivered during natural pauses in the existing soundtrack, preserving its length, in line with AD convention. This often leaves little room for describing salient visual elements, many of which remain inaccessible. Furthermore, accessing visual information may also be challenging for other audiences, including neurodiverse individuals. Neurodiversity could result in certain concepts and meanings not being recognised in the same way by everyone. Assistive technological solutions, such as AD, could make the implicit meaning associated with visual clues accessible. Recent studies have explored application of audiovisual accessibility solutions to assist cognitively diverse audiences with emotion recognition through AD-style commentary (Starr and Braun, 2020) and to aid comprehension by means of explanatory additions to audiovisual content (Deleanu, forthcoming).

The present study contributes to this research by exploring Modular AD – a novel method of creating and delivering AD which harnesses the personalisation potential of video streaming platforms, where users are not confined to a viewing schedule or length of time spent consuming streamed content. This flexibility can be exploited to accommodate new practical solutions to overcome the current restricted format of AD. Building on the existing concept of extended AD, created by automatically pausing a video to allow more time for description, the proposed method aims to systematically extend and modularise descriptions. Extended descriptions will form optional modules to enable delivery of relevant supplementary information about a visual scene along with a core description. This method facilitates a more flexible delivery of description, opening the possibility of personalising AD.

The Modular AD concept combines traditional and extended AD formats. Its potential to personalise the way users consume described streamed media content requires investigation. The knowledge that could contribute to exploring the viability of Modular AD as a new practice is distributed between different agents (Greco, 2019) – audio

describers as creators, advocates of the existing practice of extended AD, and users of AD. Therefore, the research employs multi-method user-oriented approaches to include all perspectives.

The presentation reports on phase one of the project, which examines how Modular AD method can meet diverse access needs. Specifically, it focuses on the theoretical frameworks applied during the AD scripting process in order to create personalised AD narratives. It will also present the conceptual model for the proposed modular method, the opportunities it offers, engagement of users from the early stages of research, and how the preliminary findings from empirical studies conducted to date inform a user reception study planned in phase two.

The project aims to provide foundations for further research into new practical solutions for assistive commentaries to address individual access needs of a wider range of audiences. It seeks to offer a path to an inclusive cost-effective AD creation workflow that meets the demand for different AD versions tailored to diverse user groups while accommodating increasing production volumes.

Funding information

This work was supported by the Arts and Humanities Research Council as part of the Techne Doctoral Training Partnership [Training Grant reference number AH/ R01275X/1]

References

Deleanu, A. (forthcoming). *Enhancing cognitive accessibility: Embedding 'Accessible Cues' in audiovisual material to support comprehension and enjoyment for audiences with diverse cognitive abilities* (Unpublished PhD thesis). University of Surrey, Guildford.

Greco, G. M. (2019). Accessibility studies: Abuses, misuses and the method of poietic design. In C. Stephanidis (Ed.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11786 LNCS* (pp. 15–27). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-30033-3_2

Starr, K., & Braun, S. (2020). Audio description 2.0: Re-versioning audiovisual accessibility to assist emotion recognition. In S. Braun & K. Starr (Eds.), *Innovation in audio description research* (1st ed., pp. 97–120). Routledge.

Olga Davis is a PhD student at the Centre for Translation Studies at the University of Surrey, UK under supervision of Prof. Sabine Braun and Dr Dimitris Asimakoulas.

Olga has an MA in Translation and Interpreting Studies from the University of Manchester (2020), where her dissertation was a practical audio description creation project. This work inspired a novel concept of modular audio description. Her research proposal was awarded the Techne scholarship by the Arts and Humanities Research Council UK in 2022.

Her background is in the audiovisual translation and localisation. She worked as Language Editor at Discovery Channel (Warner Bros Discovery, Inc.) between 2010-2016 and previously as freelance translator/editor of audio and multimedia guides for major museums in the UK, Europe, and the US at Antenna International.

She is a member of Audio Description Association UK. Her research interests are in audio description, access services, inclusive design, and process optimisation.

Enhanced Audio Description (EAD): making documentaries accessible through sound design

Mariana J. López, Gavin Kearney
mariana.lopez@york.ac.uk
University of York

Enhanced Audio Description (EAD) is an alternative to Audio Description (AD) in which sound design is used as a vehicle to provide accessibility for visually impaired film and television audiences, while questioning the current accessibility paradigms and their over-reliance on the spoken word to provide access. EAD is a combination of three main methods: 1) the use of sound effects to clarify the meaning of actions and settings on screen; 2) the application of spatial audio by panning sounding objects and dialogue lines to match the position of objects and actors/contributors on screen; 3) the use of first-person description/narration to provide explanations and clarifications.

EAD is an accessibility format that relies on its integration to the film and televisión workflows, and the close collaboration between accessibility experts and filmmakers. EAD aims to provide a seamless experience for audiences, while accommodating a variety of types of production by working with filmmakers to best adapt the methods to their own work.

This presentation will focus on the work we have conducted in collaboration with documentary filmmakers, reflecting on the workflows followed and how the methods were adapted to suit each production and bearing in mind the need to respect the style of the creative content and the original sound design strategy. Examples will be drawn from live action documentaries as well as animated ones, while also exploring how we have worked to convey sensitive subject matters. In this presentation we will also discuss how we worked with visually impaired people to evaluate each production and what audiences' perceptions of the EAD methods were in connection to providing an accessible and engaging experience.

Mariana J López is a Professor in Sound Production and Post Production at the University of York (UK). She is the Principal Investigator for the Arts and Humanities Research Council (AHRC) funded project 'Enhancing Audio Description II: implementing accessible, personalised and inclusive film and television experiences for visually impaired audiences', a £1m project grant which started in November 2021. In addition to her work as a researcher, Mariana teaches modules on sound design as well as accessibility to students in a variety of creative industries degrees. She is also an active sound designer. In addition to her work on accessibility, Mariana also works

in the field of acoustical heritage and historical soundscapes, having recently completed a book titled ‘Sonic Pasts: acoustical heritage and historical soundscapes’ to be published by Routledge.

Gavin Kearney is a Professor of Audio Engineering at the School of Physics, Engineering and Technology at the University of York. He is an active researcher, technologist and sound designer for immersive technologies and has published over a hundred articles and patents relating to immersive audio. He obtained his PhD in audio signal processing from Trinity College Dublin in 2010 and joined the University of York in 2011. He leads a team of researchers at York Audiolab which focuses on different facets of immersive and interactive audio, including spatial audio and surround sound, real-time audio signal processing, Ambisonics and spherical acoustics, game audio / audio for virtual and augmented reality and recording and audio post-production technique development. He is also Director of CoSTAR LiveLAB, a unique R&D facility dedicated to innovation in live performance technologies.

Storytelling: AD in the past, present and future

Marie Campbell
marie.campbell@redbeemedia.com
Red Bee Media

Great content is about great storytelling. At its heart, in providing access to the visuals of that content, audio description is, in its own way, about storytelling, too. Framed in three acts, this presentation will explore broadcast AD in the UK as one page in AD's own rich history – one version of its past, present and future.

Act 1 will be an AD origins story, as the UK celebrates 25 years of broadcast AD on pre-recorded content. The service has evolved, moving from linear to digital technologies, but the way describers tell visual stories and describe programmes hasn't, in essence, radically changed in 25 years. Describers still look for the important elements of the storytelling and describe them, distilling language to weave into the programme's audio in a way that is present but not obtrusive.

For Act 2, the Present, we will investigate the explosion on the UK accessibility stage of LIVE Audio Description for broadcast and see how this innovation both differs from, and draws on, the AD of Act 1.

If pre-recorded AD relies on script writers to tell the story of the content, in Live AD there often IS no script. In 2024, our live describers worked on content as diverse as charity telethons, celebrity dance competitions, Olympic and Paralympic ceremonies and sport – and almost all of it described off the cuff during a live broadcast. Unlike pre-recorded AD, live describers can't know when someone is going to start or stop talking in a live programme. They have the challenge of gauging *what* needs to be described and *when* to describe it. To achieve this, they rely on pre-recorded describer "muscle memory" and a knowledge of the content to tell the story.

Live AD also changes the audience experience. Pre-recorded describers try not to overlap with programme dialogue but live describers will, so where is this new editorial line, where providing access for the audience trumps what's being said, and heard, in the programme itself, and how does its position change relative to what's being described?

Finally, in Act 3, the Future, we'll explore the acceleration towards AI and automation in our industry. With technology providing at scale, offering *more* accessible content for which there is a clear and ever-growing demand, how do we harness that power while ensuring both high quality accessibility and human agency in the process? We will

look at synthetic voices and automatic script creation to see what storytelling looks like through these lenses and ask, for this story of AD, if the great American playwright Eugene O’Neill was correct when he said, “The past is the present isn’t it? It’s the future, too.”

Marie Campbell has been describing for over 20 years and is the AD Excellence Lead in Red Bee Media’s Global Innovation team. She helped set up the BBC’s first ever AD department, developing editorial guidelines and software while integrating the service into the wider broadcast chain. Marie's described for the majority of UK broadcasters and trained new audio describers in the UK, USA and Australia, tailoring AD to different English-speaking countries, while integrating workflows, maximising collaboration and providing ongoing support to new teams. In 2023/24, Marie was the Operational Lead for live audio description, implementing a brand-new service for two major UK broadcasters.

Audio Description Quality Standards, 20 years on: Insights from expert interviews with German-speaking B/VIP users

Tamara Domínguez Ebitsch
tdomingu@uni-mainz.de
Johannes Gutenberg University Mainz

As Germany marks the 20th anniversary of its first audio description (AD) guidelines (Benecke & Dosch, 2004), this presentation will revisit the concept of quality in AD, particularly considering the increasing integration of technological advances like artificial intelligence (AI) and synthetic voicing. Through in-depth qualitative expert interviews with 20 blind and visually impaired (B/VIP) German-speaking users aged 20 to 83 years old, this study, which is part of a broader PhD project, explores their previous experiences, preferences as well as views and assessments regarding the future of AD.

The interviews were structured into three key thematic areas: User Experience, User Preferences and Future Perspectives. Consistent with findings from previous studies (Fernández-Torné & Matamala, 2015; Walczak & Fryer, 2018), participants' responses suggest a general acceptance towards the use of synthetic voices, while the prospect of AI-generated AD scripts and experimental techniques such as integrated and enhanced AD were met with caution. Participants emphasized the need for transparency and user involvement in the potential development of AI-driven AD, advocating for user testing and clear labelling whenever synthetic voices or AI content are used. Additionally, some participants highlighted concerns about the recent influx of untrained AD practitioners, attributing a perceived decline in AD quality to the rapid expansion of AD services without corresponding increases in professional training and unified standards. The interviews also touched on topics long discussed in AD reception research, such as the balance between subjectivity and objectivity in descriptions (Schaeffer-Lacroix, Reviers & Di Giovanni, 2023) and the desirable degree of syntactic complexity (Mazur, 2022).

This research underscores the necessity of revisiting Germany's AD quality standards to adopt a more user-centred approach, ensuring that future guidelines and practices are both reflective of and responsive to the needs of blind and visually impaired audiences in light of technological and professional developments in the field.

References

Dosch, E., & Benecke, B. (2004). Wenn aus Bildern Worte werden: Durch Audio-Description zum Hörfilm. Bayerischer Rundfunk.

Mazur, I. (2022). Linguistic and textual aspects of audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 93–106). Routledge.

Fernández-Torné, A., & Matamala, A. (2015). Text-to-speech vs. human voiced audio descriptions: A reception study in films dubbed into Catalan. *The Journal of Specialised Translation, 24*, 61–88.

Schaeffer-Lacroix, E., Reviers, N., & Di Giovanni, E. (2023). Beyond objectivity in audio description: New practices and perspectives. *Journal of Audiovisual Translation, 6*(2), 1–7.

Tamara Domínguez Ebitsch is a doctoral candidate and lecturer at the Faculty of Translation, Linguistics and Cultural Studies at the University of Mainz. She holds an M.A. in Translation from the same department and a B.A. in Translation and Interpreting from the University of Granada. Her research interests include qualitative AD studies, AD reception in the theatre and media accessibility in general.

Live Audio Description During a Film Shooting

Elena SV Flys, Antonio Villalba
elena.sanchezvizcaino@taiarts.com
TAI University Center for the Arts

When Udo and Fels (2010a) emphasized the importance of designing accessible tools from the outset, grounded in universal design principles, they also mentioned the role of aesthetics. The problem remains when audio description is integrated into the artistic project only after its completion. What would be the outcome if the audio describer became one more agent of the creative process? In this paper, we explore the incorporation of the audio describer into the filmmaking process through an experimental practice. Film projects can be described as a combination of multiple scenes (set, sound design, voice over, etc.), and in this project, we sought to reclaim the audio descriptor's scene. To achieve this, we engaged in a live experimental process grounded in the improvisational collaboration between all participants (musicians, camera operator, actors, director, and VFX technician).

In the field of audio description, there has been ongoing debate regarding whether it should be developed with strict objectivity (e.g. Snyder, 2008), or whether it should align the aesthetics of the piece and the audio describer 's interpretive role (Candlin 2003; Cox, 2017; Fryer, 2017; Kleege 2018; Szakowska, 2013; Udo & Fels 2010a). Thus, in this work we embrace the value of personal interpretation, drawing on prior experiences such as Katsudo Benshi's film narration (Burch, 1979; Dym, 2000), to illustrate how the audio describer can emerge as a creative agent within the filmmaking process. Using an ethnographic methodological approach, both the director and the audio describer document the outcomes of this process, analyzing the implications for accessibility and the creative objectives of experimental filmmaking, as evidenced by the results of filming PUGNA, an experimental film practice where the audio describer and the descriptions themselves played a crucial role in shaping the creative process for all involved. Through observations from all agents and focus groups amongst the team, preliminary results show that the understanding of the visual components from an inclusive perspective (by means of audio description) influences the whole creation process. For example, it has promoted the initiative of exploring the image creation of the film through depth cameras which represent the relief of an image instead of the superficial layer. This demonstrates how the inclusion of accessibility design affects the creators and their conceptualization of their own artcraft, such as the search for a more multifaceted sensorial experience provoking a "shock of thought" (Thompson, 2009, p.130). Thus, rethinking the strange to become knowledge (Ahmed, 2000, p.55). This is also

reflected in the sound design which searches for a more tactile sound space. All of this show how this conceptualization of inclusiveness from the very beginning affects the whole spectrum of creation, opening new audiovisual languages and forms of creation. This impact will also be measured by inviting audience members with and without disabilities to encounter the final outcome and share their experience.

Dr. Elena SV Flys is the head of the Research Department at TAI University Center for the Arts in Madrid where she teaches arts administration and research techniques among others. Her research focuses on accessibility to the arts, audience reception, co-creation, and social integration. Around these topics, SV Flys has presented at multiple national and international conferences, has diverse publications on the subjects, and has received several research projects. She designs accessibility for theatre and music productions both in the US and Spain. Her two latest publications are "Artistic co-creation: how art students view co-creation and how it could be integrated in the arts curriculum" for Arts and Humanities in Higher Education and "Improving Accessibility and Inclusion in Performing Arts For People with Disabilities: Moving Beyond a "Checking- the boxes" Approach" in the The Oxford Handbook of Arts and Cultural Management.
Orcid: 0000-0003-3124-9362

Antonio M. Villalba is currently working on his doctoral studies at the CINEMA group of the Universitat Pompeu Fabra (Barcelona), where he also completed a Master's degree in Film Studies. His research focuses on film acting, both from an ethical and technical point of view. He has created a laboratory of acting experimentation called LICET. He also teaches History of Cinema at Escuela TAI (Madrid), where he works in the research department. In addition, he develops his career in the industrial sector as a production manager, his latest works being the short films 'Se acabó el verano' (2023) and 'La nana de la bruja' (2023); and also as a director, his latest works being the short films 'Los gritos' (2023) and 'Tiempo 4' (2023).

AD PITCHES I

Co-designing artwork audio descriptions in Australia

Sarah Barron, Janice Rieger
sarah.barron@griffith.edu.au | j.rieger@griffith.edu.au
Griffith University, Australia

This presentation advocates for the use of co-design to develop artwork audio descriptions (AD) through sharing insights from case-studies in Australian art galleries and museums. In Australia, the implementation of artwork AD in galleries and museums is currently low, and unlike all other English-speaking countries in the OECD, there are no legislative requirements nor industry standards for AD (Seeley, 2022). This presents a need to build awareness and understanding of artwork AD in Australia, to provide equitable access to art exhibitions, particularly for people who are blind or have low vision. Co-design provides an inclusive framework to produce accurate and engaging AD by working with individuals who are most affected by the outcomes and who have the required expertise (Hadley & Rieger, 2021; Rieger et al., 2023). This model focuses on 'designing with, not for', and involves people with expertise through their lived experience and/or professional practice, working together to share their knowledge and learnings (McKercher, 2020, p. 8). Despite the benefits of co-design and that engaging with people who are blind or have low vision strengthens the outcomes and avoids assumptions, this approach to AD is not practiced by most audio describers in an Australian and European context (Hadley & Rieger, 2021; Hutchinson & Eardley, 2019; Perego, 2022).

Through art and design practice-led research as part of Barron's doctoral studies at Griffith University, and Rieger's nationally funded Australian Research Council project, *Co-creating Cultures of Inclusion: Redefining Access to Cultural Heritage* (2023 – 2026), we are investigating how Australian art galleries and museums can co-design AD. This project recognises existing research into the benefits of AD for all as an inclusive rather than segregated offering, and the potential for AD as an integral part of the continual and critical evolution of galleries, museums and creative practice (Chottin & Thompson, 2021; Eardley et al., 2022; Fineman & Cock, 2022; Hutchinson, 2019; Hutchinson & Eardley, 2019; Soler Gallego, 2022; Taylor & Perego, 2022). We began co-designing artwork AD in 2019, developing our understanding and drawing on our professional expertise and studies in the fields of art and design. We work with Sarah Empey, an access and inclusion consultant who is legally blind, as well as a broader project-informed network including people in the blind and low vision community, artists, curators, registrars and

designers. This presentation aims to contribute to contemporary understandings and practices of co-designed artwork AD through sharing insights from recent case-studies of our work in Australian art galleries and museums.

References

Chottin, M., & Thompson, H. (2021). 'Blindness gain' as worldmaking: Audio description as a new *partage du sensible*. *Esprit Créateur, 61*(4), 32–44. https://doi.org/10.1353/esp.2021.0045

Eardley, A. F., Thompson, H., Fineman, A., Hutchinson, R., Bywood, L., & Cock, M. (2022). Devisualizing the museum: From access to inclusion. *Journal of Museum Education, 47*(2), 150–165. https://doi.org/10.1080/10598650.2022.2077067

Fineman, A., & Cock, M. (2022). Audio description in museums: A service provider perspective. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (1st ed., pp. 215–231). Routledge. https://doi.org/10.4324/9781003003052

Hadley, B., & Rieger, J. (2021). Co-designing choice: Objectivity, aesthetics and agency in audio-description. *Museum Management and Curatorship, 36*, 1–15. https://doi.org/10.1080/09647775.2021.1878469

Hutchinson, R. (2019). *Museums for all: Towards engaging, memorable museum experiences through inclusive audio description* [Doctoral dissertation, University of Westminster]. https://doi.org/10.34737/v01vw

Hutchinson, R., & Eardley, A. (2019). Museum audio description: The problem of textual fidelity. *Perspectives, 27*(1), 42–57. https://doi.org/10.1080/0907676X.2018.1473451

McKercher, K. A. (2020). *Beyond sticky notes: Co-design for real: Mindsets, methods and movements* (1st ed.). Beyond Sticky Notes.

Perego, E. (2022). The audio description professional: A sociological overview and new training perspectives. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (1st ed., pp. 265–279). Routledge. https://doi.org/10.4324/9781003003052

Rieger, J., Hadley, B., Barron, S., Boulton, S., & Parker, C. (2023). Codesigning access: A new approach to cultures of inclusion in museums and galleries. In A. Cachia (Ed.), *Curating access: Disability art activism and creative accommodation*. Routledge.

Seeley, K. (2022). Audio description in Australia. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (1st ed., pp. 560–578). Routledge. https://doi.org/10.4324/9781003003052

Soler Gallego, S. (2022). (Re)Imagining the museum: Communicative and social features of verbal description in art museums. *Disability Studies Quarterly, 42*(1), Article 1. https://doi.org/10.18061/dsq.v42i1.7287

Taylor, C., & Perego, E. (Eds.). (2022). *The Routledge handbook of audio description* (1st ed.). Routledge. https://doi.org/10.4324/9781003003052

Sarah Barron is a PhD student and research assistant at Griffith University, Australia. Her research project focuses on co-designed audio descriptions in art galleries and museums, particularly for people who are blind or have low vision She is also an arts worker with over 15 years of experience in the visual arts sector. Her current roles include Venue Support Officer at Pine Rivers Art Gallery and committee member of the Australian Museums and Galleries Association, Queensland branch (AMaGAQ). In addition, she engages in freelance arts and accessibility work, primarily through co-designing artwork audio descriptions. Sarah is passionate about accessibility, inclusion and engagement.

Dr Janice Rieger has worked across the cultural sector in Australia, Canada and the USA since 1996. Janice is currently an Australia Research Council Fellow (ARC) conducting research for Co-creating Cultures of Inclusion: Redefining Access to Cultural Heritage, a world-first study to create an innovative co-design model of practice and education for the GLAM sector. She is also working with Creative Australia and Arts Access to create the first national disability arts archive through an Australian Research Council grant. Janice is committed to increasing access across the cultural sector in Australia and internationally.

Audio description of one of the world's largest street parades: translating Brazilian *Carnaval* into words

Ana Julia Perrotti-Garcia
drajulia@gmail.com
Pontifical Catholic University of Minas Gerais, PUCMG, Brazil

Audio description (AD) is a kind of audiovisual translation where non-verbal language is translated into verbal language. As it happens in other types of translation, there are cultural and terminological aspects that impact the process of converting visual information into a verbal text.

The aim of this paper is to present the challenges and joys of translating images into words for a group of visually impaired people during the Brazilian Carnaval.

Carnaval in São Paulo is one of the world's largest festivals. Its parades gather samba schools which perform in the Sambadrome for three whole nights. People dress up in colorful costumes, decorated with feathers, fur, sequins, glitter, mirrors, beads, metal and precious stones. In 2023, the three nights of the festival included twenty-two parades, each featuring one of a varied and diversified set of themes, ranging from African religion to tropical birds and animals, traditional cuisine, and even famous Brazilian singers. Each member of the AD team had to write audio description scripts for eight samba schools. The strategies described below were used for all schools, focusing on their specificities.

Script preparation involved an in-depth study of the parade's topic, based on the information provided (200-500-page technical documents for each school) and research of cultural elements. After their preparation, the audio description scripts were submitted to the AD advisor and revised based on their suggestions.

Technical visits to the Sambadrome were also made, providing insights for descriptions of the grandstands and parade floats.

During the parade, the AD team worked in pairs, with a rotation of the active pair at the end of each performance to allow time for physical and mental recovery.

After the parade, we talked to people in the audience who had watched the performances with audio descriptions (both visually impaired people and their family and friends). Positive aspects highlighted by them included increased understanding of the show, greater absorption of cultural content, and access to more visual information present in the performances.

One of the specific difficulties the team faced was describing imaginary beings made real during the parade, like a three-headed dragon with sparkling colors making rhythmic movements, an eagle with a 20-meter wingspan crossing the avenue where the parade took place, or Saint-Exupéry's Little Prince, just to name some. Each parade also included several allegorical cars (huge vehicles designed specifically for the parade and decorated according to its theme). On the parade floats, some participants also danced, performed, interacted with each other and, sometimes, with the audience. For the description of these visual elements, we used strategies drawn from audio description for theater.

This experience made it clear to us that, for live events of this magnitude, it is essential to work on teams whose members are in tune with each other and ready to conduct in-depth research and detailed analyses. Translating all these images into words, without losing the art and beauty of an event of such huge proportions, was certainly the job of a lifetime.

Ana Julia Perrotti-Garcia is a professor of Audio description and Medical Translation, she holds a PhD in English Language and Literature (USP- Brazil), and a master's in Applied Linguistics (Pontifical Catholic University SP), specialist in Audio description (PUCMG). Mentor of Students from the Master's Degree in Translation for International Communication Program, Universidade de Vigo, Spain (2024). Associate Professor, Universidade Autónoma de Lisboa (UAL), Portugal (2024). Among her research topics, she studies translation and adaptation, audio description and accessibility. She has been working with audio description since 2009, in live concerts, conferences and also in movies, photographs and works of art. Ana Julia is the coorganizer/author of the book: Live and In Color: Case Reports of the audio description of live events (Amazon) and CEO of TraduSound Audio-Visual Accessibility. She has delivered lectures and courses in Brazil, United States, England, Ireland, Australia, Portugal, among others.

The voice of the Valencian ninots: A reflection on neutrality

Vanessa Gumier García
gumier@uji.es
Universitat Jaume I, Spain

Ninots - caricature pieces - are the principal element of monuments erected on the streets as part of a festivity celebrated in spring in some towns of the Valencian Community in Spain (UNESCO, 2016). Article 27 of the Universal Declaration of Human Rights (United Nations General Assembly, 1948) prescribes equal access to culture for all, hence the need to study access for the blind and partially sighted to the Valencian ninots. Given the sculptural nature of a ninot, its audio description takes the form of an audio guide in the same way as artworks exhibited in museums. Although neutrality remains latent in museum audio description guidelines, more marked approaches keep being explored (Soler Gallego, 2019; Bartolini, 2023) not only in terms of linguistic content but also voicing (Matamala & Machuca, 2022).

As part of an observation of access at Valencian ninots, this paper delves into the correlation between voicing and neutrality (Fryer, 2016; AENOR, 2005) with the focus on voice source, that is, the use of synthetic and human voices in the distribution phase. Emotions are conveyed in speech through certain prosodic features, like pitch, volume and speed rate (Frick, 1985). When asking users (Matamala & Machuca, 2022), other variables such as accent, articulation, gender, style or AD-related features should also be added to the list. An acoustic analysis based on parameters by Matamala & Machuca (2022) is conducted on a corpus consisting of four Spanish and Catalan audio guides - two delivered by a synthetic voice and two delivered by a human voice - available in 2024. The latter are either generated by professionals or festivity stakeholders. Parameters analysed include emotion, pitch and inflection, accent, articulation and pronunciation, volume and speed rate, gender, style and AD-related features. The results would be applicable not only to access to Valencian ninots but also to access to museums and similar festivities.

References

Bartolini, C. (2023). Museum audio descriptions vs. general audio guides: Describing or interpreting cultural heritage? *Journal of Audiovisual Translation, 6*(2), 77–98. https://doi.org/10.47476/jat.v6i2.2023.253

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*(3), 412–429.

Matamala, A., & Machuca, M. (2022). Neutral voices in audio descriptions: What does it mean? *Babel Revue Internationale de La Traduction / International Journal of Translation, 68*(5).

Soler Gallego, S. (2019). Defining subjectivity in visual art audio description. *Meta, 64*(3), 708–733. https://doi.org/10.7202/1070536ar

UNE 153020. (2005). *Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías.* Asociación Española de Normalización y Certificación.

UNESCO. (2016). Valencia Fallas festivity. UNESCO. Retrieved from https://ich.unesco.org/en/RL/valencia-fallas-festivity-00859?RL=00859

United Nations General Assembly. (1948). *Universal Declaration of Human Rights.*

Vanessa Gumier García is a PhD student in Applied Languages, Literature and Translation at Universitat Jaume I, supervised by Dr Irene de Higes Andino. She is a graduate in Translation and Interpreting from Universitat Jaume I, where she has also completed a Master's degree in Secondary Education, Vocational Training and Language Teaching. She has collaborated with the research group Translation and Communication in Audiovisual Media (TRAMA) through the Estudia i Investiga programme at Universitat Jaume I. She has also worked as a freelance translator for dubbing studios and film festivals. In November 2023, she joined the research group TRAMA within the Department of Translation and Communication at Universitat Jaume I as a predoctoral scholar (ACIF/2022/246). She has just undertaken a three-month-long research stay at the University of Antwerp under the supervision of Dr Nina Reviers. Her thesis project explores accessibility for blind and visually impaired people at Valencian Fallas.

Testing Clipchamp as an Automated Text-to-Speech Recognition Software for Audio Description Research

José-Luis Castillo-Flores
luis.castillo@ulpgc.es
Universidad de Las Palmas de Gran Canaria

Up until now, the scarce academic literature on audio description (AD) tools revolves around programs that have been developed for that very specific purpose (Fryer, 2016; Minutella, 2022). Steadily more, regular video editors offer options that could come in handy for the multifaceted nature of AD research: from cutting a video to voicing a target text (TT). Clipchamp, a free web-based user-friendly video editing software, is an example of this. Among its multiple features there is one that could be of particular interest for scholars conducting research on AD with no financial support, namely the automated text-to-speech recognition feature. Other sound editing options include volume adjustment for individual tracks (e.g. music vs audio description), sound fading or background noise suppression, plus an array of AI-generated natural sounding voices to choose from.

In this presentation, we will share our experience operating Clipchamp to create a fully functional audio described TT that was ultimately used to test different aspects about audio description with a blind and partially sighted (BPS) audience as part of a larger study. In particular, we will focus on this automated text-to-speech recognition feature, and comment on its strengths including decimal speed-adjustment or pitch control (type of voice, type of emotion, vocal pitch), and weaknesses like pronunciation mistakes in Spanish for specific chains of words or intonation flops for long segments.

The text-to-speech recognition feature was applied to eight different clips audio described in Spanish that lasted a total of 45 minutes. These clips were then played to our BPS audience, who did not know anything about how the AD had been voiced. Immediately afterwards, they answered a Likert scale questionnaire that covered multiple aspects related to speech delivery, such as voice quality, intonation, speed, pauses, naturalness, pleasantness and overall satisfaction. The closed questions were followed by a group discussion that allowed BPS users to shape their answers and exchange impressions. Preliminary results show that our BPS audience was highly satisfied with the automated text-to-speech recognition feature, which proves that Clipchamp meets their standards and thus could be considered as a suitable software for those who are conducting AD research with no financial support.

José Luis Castillo-Flores holds a BA in Translation and Interpreting by the University of Las Palmas de Gran Canaria (ULPGC). He has completed different courses on audio visual translation at the Universitat Autónoma de Barcelona and holds an MA in Arts by the University of Hawai'i at Manoa. During his stay in Hawai'i, he worked as a lecturer and taught different classes on Translation and Spanish. He was also an advisory board member of the Center for Interpretation and Translation Studies. At the ULPGC, he has taught Subtitling at Master's level and, more recently, Documentation Applied to Translation. He is currently a PhD student at this same university and is interested in cinematic audio description.

Exploring the audio description of humour: A multimodal corpus study in the Chinese context

Zi Ye
zi.ye@bristol.ac.uk
University of Bristol

Humour is a significant element of audio-visual media used to entertain audiences. Between 1998 and 2017, humour-based comedy accounted for 28.4% of all films made by the global filmmaking industry (Stephen Follows 2018). However, humour is a challenging issue for audio description (AD). The multifaceted nature of visual humour is difficult to describe in only a few words, and its strong emotional demands contradict objectivity-centred AD methods. Few studies have examined the AD of humour (e.g. Martínez-Sierra 2010; Chorão 2021); these have tended to discuss whether AD maintains or excludes the humour without addressing more specific resolutions and have primarily considered the Western context. This study aims to build a multimodal corpus of comedy films audio described in Chinese, examine the complicated transfer process involved in translating visual humour to AD and explore the current norms associated with humour AD in China. The study focuses on the under-explored Chinese context, where humour appreciation and AD production methods differ from those in the West. Moreover, the specifics of the Chinese language, including the use of *chengyu*, a form of idiomatic expression mainly using four characters, can bring distinctive qualities to humour AD.

This presentation focuses on the annotation process and results of the multimodal corpus. Six comedy films audio described in Chinese have been incorporated into the corpus using Elan software, an annotation tool for video materials. Integrating humour theories (Yus 2012), a taxonomy of subtitling strategies (Pai 2017) and specific AD techniques (Bardini 2020), a comprehensive annotation scheme was developed, which was used to dissect the AD treatments of humorous scenes. The scheme covered multimodal aspects, but mainly cared about visual humour in three dimensions (information, culture and structure) and AD strategies at the macro- and micro-levels. Quantitative analysis was undertaken to identify the features of this multimodal corpus, with qualitative examples used to provide in-depth illustrations. The results revealed patterns of different types of visual humour and their corresponding AD solutions and the preliminary relationships between the two variables. This study will support audio describers to treat humorous elements in a more targeted manner and adopt different AD strategies for particular situations. It will also provide a labelling model for future research to support the identification of AD approaches in humour-based visual media.

References

Bardini, F. (2020). Audio description and the translation of film language into words. *Ilha do Desterro, 73*(1), 273–296.

Chorão, G. B. (2021, January 3). *The audio description of humour* [Video]. YouTube. Available at: https://www.youtube.com/watch?v=6DfWQbC6FpY&t=45s (Accessed August 15, 2024).

Martínez-Sierra, J. J. (2010). Approaching the audio description of humour. *Entreculturas, 2*, 87–103.

Pai, A. F. S. (2017). *Subtitling humour from the perspective of relevance theory: The Office in traditional Chinese* (Doctoral dissertation). University College London. Available at: https://discovery.ucl.ac.uk/id/eprint/1560820/ (Accessed August 20, 2024).

Follows, S. (2018). Genre trends in global film production. Available at: https://stephenfollows.com/genre-trends-global-film-production/ (Accessed August 17, 2024).

Yus, R. F. (2012). Strategies and effects in humorous discourse: The case of jokes. In B. Eizaga-Rebollar (Ed.), *Studies in linguistics and cognition* (pp. 271–296). Berlin: Peter Lang.

Zi Ye is a PhD student in Translation Studies at the University of Bristol and Cardiff University, funded by the South, West, and Wales Doctoral Training Partnership. Her research interests lie in audiovisual translation and media accessibility, with a focus on audio description and humour translation.
She has worked as an AD volunteer since 2021, writing AD scripts for Chinese audiovisual products and helping organise audio-described film screening events. She was a trainee in the 2021 Training Workshop for Audio Description and Subtitles for the Deaf and Hard of Hearing, held by the Entertainment Accessibility Project Team, and a trainee in the August 2023 Training Courses for Assisting the Visually Impaired People, held by the Xinmu Cinema.

PANEL 4. AUDIO DESCRIPTION RESEARCH

Mind the Boundaries: Neurocognitive and AI-Driven Insights into Event Perception in Audio-Described Films

Roger Johansson, Maja Rudling, Johan Mårtensson, Jana Holsanova
roger.johansson@psy.lu.se
Luns University, Sweden

A fundamental aspect of the human mind is its ability to perceive, understand, and remember life experiences, which we process as distinct and meaningful sequences of events, each with a clear beginning and end. This includes experiences from audiovisual media. However, while movies do not explicitly signal when one event ends and another begins, viewers still consistently and automatically perceive event boundaries. To assess the effectiveness of Audio Description (AD) in enhancing the film experience for visually impaired viewers and to understand the challenges and opportunities for automated, computer-generated video descriptions, it is crucial to explore how AD influences the perception and segmentation of a film's narrative structure.

In a recent study (Johansson et al., 2024), we investigated how visually impaired participants perceived event boundaries in a film compared to their sighted peers. The sighted participants viewed the original film, while the visually impaired participants experienced two versions of AD of the same film: one with explicit indications of key event boundaries and another where these boundaries were more subtly conveyed. All participants were asked to segment the narrative into distinct and meaningful events.

Our findings reveal that visually impaired participants perceived event boundaries similarly to their sighted counterparts, indicating that AD can effectively convey the unfolding event structure. However, when event boundaries in AD were more subtly conveyed, they were less likely to be recognized. This underscores the importance of presenting clear and explicit event boundaries in AD to enhance the cinematic experience for visually impaired viewers.

Building on these findings, we are currently conducting a brain imaging study to further investigate how event structures are experienced, understood, and remembered in films with AD. This study targets the neural activation patterns in people with and without visual impairments as they experience event boundaries in film narratives, with and without AD.
While our research offers valuable insights for improving AD practices, it also holds significant implications for the development of

automated, computer-generated video descriptions (Braun, Starr & Laaksonen, 2020; Starr, Braun & Delfani, 2020). A deeper understanding of how visually impaired users perceive, understand, and remember event boundaries in AD can inform the design of algorithms capable of accurately tracking narrative elements across time and space. Such advancements could address key challenges in maintaining continuity and coherence in AI-generated video descriptions. For instance, when applying current commercial AI models, such as ChatGPT 4.0, to identify event boundaries in the films from our studies, these models mainly identified visual contrasts at scene transitions or clear changes in visual flow. This approach contrasts sharply with human event segmentation, which fundamentally relies on a contextual understanding of the unfolding narrative, including spatiotemporal relationships among characters, as well as their intentions and goals toward each other – a capability that remains largely absent in the interpretation of audiovisual media by commercially available AI.

Roger Johansson, PhD, is an Associate Professor of Psychology and Senior Lecturer at the Department of Psychology, Lund University. His research centers on cognition, communication, and learning, with a particular focus on the relationships between eye movements, mental imagery, episodic memory, attention, event cognition, and narrative processing. Together with Jana Holsanova, he has headed interdisciplinary research projects over the past five years, exploring the psychological and neurocognitive mechanisms that underpin effective communication between sighted and blind individuals. These projects, with a special emphasis on audio description, aim to deepen our understanding of how fundamental principles of memory, mental imagery and event representations contribute to accessible communication.

Maja Rudling has a PhD in Psychology and is a licensed speech-language pathologist working as a post-doc in cognitive neuroscience at Lund university. Her research has focused on the intersection of language, communication, perception, and attention. In her post-doc she works with Jana Holsanova, Roger Johansson, and Johan Mårtensson on a research project aimed at understanding neurocognitive underpinnings of event perception in sighted and blind individuals.

Johan Mårtensson, PhD in psychology, Associate professor in cognitive neuroscience is division head of Logopedics, Phoniatrics and Audiology at the Department of Clinical Sciences Lund (IKVL). His work is predominantly interdisciplinary, combining measures of cognition and brain imaging with a focus on learning. Johan has previously been active in the steering group for CCL (one of Lund's Linné environments) and is currently active in the Birgit Rausing

Centre for Medical Humanities, a 30-year endeavor to change the medical educations in collaboration with socials sciences and the humanities.

Jana Holsanova, PhD., is Associate Professor in Cognitive Science at Lund University. Her research focuses on multimodality, cognition and communication, in particular on how we perceive images, how the interaction between language and images works and how we create mental images. Jana Holsanova is the Director of the AUDEA research group at the Cognitive Science Department at Lund University. Together with Roger Johansson, she has headed interdisciplinary research projects over the past five years, exploring the psychological and neurocognitive mechanisms that underpin effective communication between sighted and blind individuals. She is the Coordinator of a large collaboration network initiative in Sweden (ADACOM), consisting of blind users, professional audio describers, educators of audio description, and representatives from government agencies and disability organisations. Jana Holsanova is also an active Board Member of the Swedish Braille Authority at Swedish Agency for Accessible Media.

Findings from a linguistic experiment suggest that multisensory language in audio description helps listeners to generate more vivid mental imagery

Jessica Beale
J.beale@westminster.ac.uk
University of Westminster

There is little research in audio description (AD) that is cognitively-orientated (Holsanova, 2022), or focussed on sensory language. However, there is great potential to develop our understanding of how AD functions for blind and partially sighted (BPS) audiences by taking a cognitive approach to the sensory language of AD. For instance, multisensory approaches to the language of AD may be more effective than typical approaches to AD (e.g. purely visual description) because language based on multiple senses could help people develop richer mental imagery (Fryer, 2016). This is supported by research in cognitive science and linguistics, in the growing body of evidence that cognition (including language comprehension) is grounded in sensory/perceptual mental simulations (Barsalou, 2008). As mental imagery is closely related to mention simulation (Speed and Majid, 2020), and related to one of the goals of AD (e.g. ADLAB, 2014), it is a useful measure for evaluating the impact of linguistic choices in AD.

This paper will outline the potential opportunities of exploring AD through a cognitive linguistic perspective, with a particular focus on sensory linguistics (Winter, 2019), and the role of embodied cognition in AD experiences. This presentation will outline the methodology and results from an online linguistic experiment testing different approaches to the AD of plants. This study is part of a wider research project exploring audio description in botanic gardens, so some findings are specific to this context, but many are applicable in AD more broadly, as well as relevant to the study of cognitive linguistics.

This study asks: is multisensory AD more effective than visual-only AD in helping people to generate vivid multisensory imagery, and how does this vary between groups? (e.g., level of vision, or age at onset of visual impairment) The study tests whether multisensory descriptions are more effective than visual-only descriptions, and whether multisensory description led by visual language or tactile language is most effective. There were 156 participants (63 sighted, 93 BPS participants – with participants grouped further based on level of visual impairment (VI) and age at onset of VI). Participants listened to nine audio description clips of different plants, and after each AD clip, rated the strength of the imagery they could generate for that plant across five modalities, using an adapted version of the Plymouth Multisensory Imagery Questionnaire (Andrade et al., 2014),

as well as answering questions on the usefulness of the description, the level of effort it took to imagine the plant and how much they enjoyed the description.

The findings suggest that multisensory description helps all people (regardless of their level of vision) to generate more vivid mental imagery across all sensory modalities, with tactile-led description enhancing tactile imagery. Perhaps most remarkable is that multisensory description helps listeners to generate more vivid visual imagery, as well as stronger imagery in all other modalities. In addition, these findings support Fryer's (2016: 172) hypothesis that description based on multiple senses might help people to generate more vivid multisensory imagery.

References

ADLAB. (2014). *ADLAB Guidelines for Audio Description.* The ADLAB Project. Available at: http://www.adlabproject.eu/Docs/adlab%20book/index.html (Accessed May 10, 2024).

Andrade, J., Panabokke, N., & Baddeley, A. (2014). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *The British Journal of Psychology, 105*(4), 547–563. Available at: https://doi.org/10.1111/bjop.12050.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645. Available at: https://doi.org/10.1146/annurev.psych.59.103006.093639.

Fryer, L. (2016). *An Introduction to Audio Description: A Practical Guide* (1st ed.). London: Routledge.

Holsanova, J. (2022). A cognitive approach to audio description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 57–77). Oxford: Routledge.

Speed, L., & Majid, A. (2020). Grounding language in the neglected senses of touch, taste, and smell. *Cognitive Neuropsychology.* Available at: https://doi.org/10.1080/02643294.2019.1623188.

Winter, B. (2019). *Sensory Linguistics: Language, Perception and Metaphor.* Amsterdam/Philadelphia: John Benjamins Publishing Company. Available at: https://doi.org/10.1075/celcr.20.

Jessica Beale is a PhD student at the University of Westminster, and her research is based at the Chelsea Physic Garden. This research is funded by Techne, and in partnership with VocalEyes, a charity which

aims to bring art and culture to life for blind and visually impaired people, and the Chelsea Physic Garden, a 350-year-old medicinal garden in the centre of London.
She is interested in the application of cognitive linguistics and the theory of embodied cognition in relation to designing and evaluating audio description. Her research focuses on the impact of language use in audio description for blind and partially sighted people, and the effect that this has on cognition, imagination and enjoyment for visitors to botanic gardens.

Audio Description Expertise and Quality: Unraveling A Conceptual Yardstick

Harun Dalli, Anna Jankowska, Iris Schrijver
harun.dalli@uantwerpen.be
University of Antwerp

What makes an expert describer and a high-quality audio description (AD)? While the demand for audio description continues to rise, our theoretical understanding of expertise and quality remain largely unexplored. The ADEQUACY project delves into this critical gap, aiming to decode the elements that systematically differentiate a good describer and description from others. Our research adopts a mixed-methods approach to uncover these understudied aspects of AD.

On the quantitative front, we integrate methods from Cognitive Translation Studies (CTS) and Translation Process Research to non-invasively monitor the scripting process of 20 describers as they describe the short film *What Happens While* (Nia, 2015), created for the purposes of the Visuals into Words project (Matamala & Villegas, 2016). The aim is to understand how describers of varying experience levels allocate the three types of translation effort identified by CTS (temporal, cognitive, and production) and how these efforts correlate with expertise—an attribute not solely defined by years of experience. The keyboard interactions of the subjects are recorded via Inputlog (Leijten & Van Waes, 2013), which serves as the main research tool to measure the use of external resources, pausing patterns, production speed, and total task duration. Since keylogging alone lacks the contextual depth to fully understand underlying cognitive processes (Hvelplund, 2011), specifically during pauses, we complement it with eye-tracking (via Tobii Pro Fusion 120hz) to explore differences in visual attention among describers of varying seniority. We borrow eye-tracking measures from expertise studies and test them on the AD process. This includes analyzing fixation patterns within predetermined areas of interest (AOIs), relative dwell time, and transition frequencies (Holmqvist et al., 2011).

On the qualitative front, we employ methods from Participatory Action Research (PAR) to engage directly with blind users, describers, and project managers to understand their perceptions of AD expertise and quality through multi-round DELPHI surveys and focus groups. The knowledge gathered from industry stakeholders and users will inform the further advancement of the AD quality assessment framework developed by Fryer (2019). The scripts produced during the experimental phase will then be evaluated by two blind AD consultants and two external reviewers according to this revised assessment criteria. This evaluation will help us identify which

performance behaviors lead to high-quality descriptions and how different levels of experience influence the scripting process. By actively involving both AD creators and users, we ensure that our approach narrows the maker-user gap (Greco, 2019) that is prevalent in media accessibility.

At the ARSAD conference, we aim to introduce our project and share our initial findings from the scripting experiment conducted in autumn 2024.

Harun Dallı is a Ph.D. fellow at the Department of Applied Linguistics, Translation, and Interpreting at the University of Antwerp, where he is working on his FWO-funded research project entitled '(ADEQUACY) Audio Description Expertise and Quality: Unraveling A Conceptual Yardstick' under the supervision of Prof. Anna Jankowska and Dr. Iris Schrijver. Previously, Harun contributed to research projects on pivot audiovisual translation at University College London and literary machine translation at Boğaziçi University, where he received his Master's degree in Translation Studies in 2023 with a thesis on pivot subtitling. His research interests include audiovisual translation, corpus linguistics, media accessibility, and the translation process. Beyond academia, Harun has professional experience in subtitling, having translated titles such as *Squid Game: The Challenge*, *DEPP V HEARD*, *j-hope IN THE BOX*, and *Raising Voices*.

Audio Descriptive Commentary in Football

Inma Pedregosa
inma.p@btinternet.com
University of Westminster

Blind and partially blind (BPB) fans attending a live football match can sometimes use audio description (AD), also known as audio-descriptive commentary (ADC) provided by the clubs. However, although football is the most popular spectator sport in the world and has such a prominent role in the lives of individual, communities and countries, there has been virtually no research about this type of AD. Unlike other forms of AD, ADC is not the only narrative (commentary) available to audiences who have limited or no visual access to a match. At a time where football TV rights are worth billions, radio commentary is a free and inclusive way for fans in Britain who cannot see the game to follow their teams and experience the live event as it happens.

This interdisciplinary thesis uses a (quantitative and qualitative) mixed-method approach to understand what the current practice of ADC in Britain is and compare it with another 'translational action' - radio commentary. It deals with some of the core preoccupations of AD research, i.e., the nature of the source text and the desirability of objectivity, and adds a feature inherent to competitive sport, i.e. partisanship. The research takes a three-pronged approach: exploring the textual content of both types of commentary, the practitioners' views and the users' opinions. Four studies were carried out as follows:

Study 1 examinedcurrent understandings of football ADC and football radio commentary, by means of a practitioners' survey. The aim of ADC and radio was revealed to be a way of bringing an emotional experience to the listener rather than an objective visual to verbal translation. Audio describers and radio commentators also showed wide agreement about the importance of the elements presented to them related to describing the football event, the type of language use and being neutral when narrating. There were only a few differences: for instance, for audio describers, describing colours and players' styles, and using football terminology was more important than for radio commentators.

Study 2 explored the content of ADC and radio commentary via textual analysis. A coding model was devised to ascertain what the potential source text might be and how it is translated. Results show that, for instance, in radio, 20% of the content is unrelated to the actual football event, whereas in ADC, the percentage is 4%. In ADC, more language was objective than interpretative whereas in radio

there was no difference. In radio, there was a higher number of semantic units denoting opinion than action and spatial units combined, whereas ADC devotes more time to describe action and spatial aspects than opinion than radio.

Study 3 carried out a textual analysis, using the same coding model as study 2, comparing the ADC that is created specifically for Blind and Partially Blind (BPB) audiences in-stadium, with commentary that is commentary being currently offered to BPB as 'access' but which is not created specifically for them as an audience (such as the one provided by clubs via their apps). Results showed some similarities between both types, for instance having the same amount of football specialised language. In terms of differences, ADC had more objective units and less interpretative units than Other Commentary, and ADC included more references to space than Other Commentary.

Study 4 used a survey methodology to examine the experiences of BPB and sighted football fans. It looked at their experiences of live matches and of radio commentary, and for BPB fans, their experiences of ADC. It buildt on the previous research to examine the nature of the source text in football, and the expectations and preferences of BPB and sighted listeners. BPB and sighted listeners showed wide agreement about the importance of the elements presented to them related to describing the football event and the expectations of commentary. There were some differences though, such as BPB fans attaching more importance to listening to a commentary from the point of view of the team they support than sighted listeners.

This research, carried out with the collaboration of football clubs and the contribution of audio describers, radio commentators, and BPB and sighted football commentary listeners, builds on the knowledge of audio description in practice and ultimately aims to inform ADC and radio commentators' practice to better serve their listeners.

Inma Pedregosa is a PhD researcher at the University of Westminster and is Visiting Lecturer at the University of Westminster and King's College London. Her overarching research interest is mediation in its different forms, including audio description, descriptive subtitles and mediation in language teaching. She has consulted OFCOM on the quality of live subtitles and more recently, FIFA on the quality of their audio descriptive commentary. Outside academia, she also works as a translator, subtitler and interpreter.

The heartbeat of creativity in audio description. A pilot heart rate study on multimodal objectivity and subjectivity

Alejandro Romero-Muñoz
alromero@uji.es
Universitat Jaume I, Spain

The appropriateness of subjectivity over objectivity (and vice versa) in audio description (AD) is an ongoing debate far from having reached a clear consensus, where two opposing perspectives, the professional and the academic one, seem to collide. Whereas some professional guidelines support the view that an objective AD is the advisable option, some scholars have pointed to the advantages of more subjective, creative, or narrative alternatives (Kruger, 2010; Fryer & Freeman, 2012; Soler Gallego & Luque Colmenero, 2023).

Meanwhile, part of AD research is making inroads into experimental methodologies, proof of which can be studies on AD and eye-tracker (Krejtz et al., 2012; Kruger, 2012; Orero & Vilaró, 2012; Di Giovanni, 2014; Mazur & Chmiel, 2016), text-to-speech AD (Szarkowska, 2011), the automatic generation of AD (Hasegawa-Johnson et al., 2017), or physiological measurements applied to AD, such as heart-rate (Jankowska et al., 2022), cortisol levels (Ramos Caro, 2015; Ramos Caro, 2016; Rojo López et al., 2021), or skin conductance response (Sterna et al., 2023), among others. In this vein, our proposal presents an experimental pilot study using a heart rate sensor to measure visually impaired participants' response to different combinations of subjective and objective multimodal elements in some AD clips. To do so, objective and subjective multimodal elements from an AD clip from Netflix's Money Heist were analyzed following Romero-Muñoz's (2023) proposal that any AD displays a certain combination of creative elements within its script. Afterwards, objective components were turned subjective (and the other way around), which resulted in two versions of Money Heist's clip in subjective-objective terms: the original and the manipulated one. Both versions were recorded by a professional studio, which eventually yielded four AD: the original AD recorded by a man, the original AD recorded by a woman, the manipulated AD recorded by a man, and the manipulated AD recorded by a woman. Participants were randomly allotted one of these four possibilities, they filled in orally a demographic survey, a presence test (ITC-SOPI), and a comprehension test followed by a semi-structured interview.

Throughout the procedure, an UPTIVO Belt-D heart rate sensor was attached to the participants' chest to measure beats per minute, average heart rate, maximum heart rate, minimum heart rate, etc. Moreover, the UPTIVO Belt-D allows the measurement of beats per minute every few seconds of a given clip, so that the fragments

(whether objective or subjective) giving rise to higher heart rate levels can be located. Results were triangulated with the tests and interviews to obtain empirical data about the combination of subjective and objective multimodal elements in AD that induces higher heart rate and presence levels, the elements that both enable better comprehension and are better considered by users. Future research following this pilot study, which should be applied to an appropriate number of participants, could have a considerable impact on the debate over subjectivity in AD, since data would provide empirical evidence about an immersive combination of objective and subjective multimodal components in AD scripts to be considered by guidelines.

Alejandro Romero-Muñoz holds a BA in Translation and Interpreting (Universitat de València) and in English Studies (UNED), a MA in Audiovisual Translation (Universitat Autònoma de Barcelona), and a MA of Research in Translation and Interpreting (Universitat Jaume I), where he is now a predoctoral researcher. His main research interests include audiovisual translation (particularly audio description), multimodality, and objectivity. He is a member of the UJI research group TRAMA (Translation for the Media and Accessibility). Beyond his academic profile, he has worked as a professional audio describer.

PANEL 5. AN ACCESSIBILITY TRIP AROUND THE GLOBE

Audio description on streaming platforms: A snapshot of the present language offer

[1]Blanca Arias-Badia, [2]Katixa Agirre-Miguelez, [3]Antía María López-Gómez, [4]Naiara Larrakoetxea-Salgado, [2]Marijo Deogracias-Horrillo, [2]Josu Amezaga-Albizu
blanca.arias@upf.edu
[1]Universitat Pompeu Fabra
[2]University of the Basque Country, UPV/EHU
[3]University of Santiago de Compostela, USC
[4]Basque Federation of Associations of Deaf People and Spanish Confederation of Deaf People

Audio described content is still scarce on present streaming platforms, especially when it comes to content offered in audio versions in languages other than English. The language offer is constantly updated and platforms do not easily disclose this type of information, which makes it difficult for both audio description researchers and end users in general to become aware of the present landscape. This paper presents results of a research project recently commissioned by the Group of The Greens/EFA, from the European Parliament (ref. EHU2022/5841). The aim of the project was to obtain an updated snapshot of the present language offer on three major platforms, namely Netflix, Amazon Prime, and Disney+, considering both source versions and translations of the platforms' titles. In our paper, we focus on the results regarding audio description: in total, 107,247 audio description tracks were found in the sample. Audio description is available in 12 % of the total content, with English showing a dominant presence in this access service. In our presentation, we will delve into the audio description offer found for each of the platforms and discuss the implications for prospective European language and accessibility policies, with a focus on the results obtained for minority languages. Our paper draws an explicit link between linguistic rights and the right to participate in cultural life by means of access services: the results of the study point to a serious accessibility issue, in which access to culturally mainstream audiovisual content is hindered. The situation may be interpreted as a discriminating factor with direct consequences on persons with vision impairments' right to partake in cultural life with equal opportunities.

Blanca Arias-Badia is a lecturer at the Department of Translation and Language Sciences at Universitat Pompeu Fabra. She is a member of the TraDiLex research group and the training coordinator of the AccessCat transfer network. She holds a PhD in Translation and Language Sciences from UPF (2017) and did a postdoc in Media

Accessibility at Universitat Autònoma de Barcelona (2018-2020). She was a visiting researcher at University College London (2014) and the University of the Basque Country (2016). She is the principal investigator of the project UnivAc, which focuses on accessibility at universities, as well as co-leader of PlanAccess, a research transfer project aimed at offering accessibility plans for the educational and cultural sectors. She is active in dissemination: she is the research coordinator of the Catalan Association for the Promotion of Accessibility (ACPA), as well as co-founder of En sincronía, a podcast devoted to audiovisual translation and accessibility.

Katixa Agirre-Miguelez is Graduate in Audiovisual Communication (UPV/EHU, 2003), Master in Cinematographic Scriptwriting (Septima Ars, 2004) and PhD in Audiovisual Communication (UPV/EHU, 2010), with a thesis that analyzed the representations of the Lolita myth in cinema. Hollywood. She was a postdoctoral researcher during the three-year period 2011-2013, doing stays at Queen Mary, University of London and University of East Anglia, where she carried out a transnational audience study on the series Mad Men. Since 2013 she has been an adjunct professor and since 2019 an associate professor of the Department of Audiovisual Communication and Advertising and, as a member of the NOR group, she researches new Basque cinema and its audiences, with special interest in streaming platforms. She is also a fiction writer in Basque, and her stories and novels have been translated into fifteen languages.

Antía María López-Gómez holds a PhD in Communication Sciences from the Complutense University of Madrid. Professor of Audiovisual Communication and Advertising at the Faculty of Communication Sciences at the University of Santiago de Compostela. She is the author of several books, including: "Políticas de comunicación e identidad cultural: estrategias gubernamentales sobre la comunicación social" or "La dimensión de lo axiológico en el discurso del spot publicitario". Co-author of the chapter with Enrique Castelló "The digital mask on social media: The tyranny of beauty filters among digital natives", in Magalhaes, Luisa & Oliveira, Candido (coords.): Digital Maks and Human connections, PalgraveMacmillan, 2023. His research focuses on textual theory and analysis and the study of European audiovisual policies.

Naiara Larrakoetxea is a deaf researcher, currently working at the Department of Political Science and Administration of the University of the Basque Country. She is currently a visiting student at Gallaudet University in the United States, and as a PhD student, her work focuses on the analysis of the political culture of deaf people, one of the key pieces of which is linguistic accessibility. As a deaf person, she has held positions of political responsibility within different

representative entities of deaf people and the field of disability, especially in relation to Language Policy.

Marijo Deogracias is a journalist on Euskadi Irratia (EITB Media), and she has just defended her thesis "Hizkuntza irisgarritasuna, ikus-entzunezkoak hizkuntza gutxituen erakusleiho" (Linguistic accessibility, audiovisual media as a guarantee of minority languages) at the UPV/EHU under the direction of Josu Amezaga. She is member of the group NOR (UPV/EHU GIC 21/053) and collaborates with ALMA - Network for research excellence in media accessibility (2020-2022), a joint project between seven research groups from different disciplines (engineering, sociology, linguistics and communication) working together on audio-visual accessibility and minoritized languages. Her research interests are focuses on minority languages (mostly, Basque language) and audiovisual media accessibility.

Josu Amezaga-Albizu is a Full Professor at the Department of Audiovisual Communication and Advertising of the University of the Basque Country (UPV/EHU). He is also head of the NOR Research Group (nortaldea.eus). He has published several works and taken part in many research projects about the relationships between language, communication and identity, specially regarding minority languages (both Immigrant and Regional Minority Languages).
He is the PI to the project Behategia (Observatory of Basque Media), where several research works are conducted about Basque language media (behategia.eus). They include research on audiences, digital analytics, automatic translation and other subjects.
He has been also the PI to ALMA - Network for research excellence in media accessibility (2020-2022), a joint project between seven research groups from different disciplines (engineering, sociology, linguistics and communication) working together on audio-visual accessibility and minoritised languages.

Assessing the Accessibility of Multimedia Content in Polish and EU Public Institutions

Marta Żaczkiewicz, Anna Jankowska
marta@deskrypcja.pl
University of Anwerp

Ensuring digital accessibility is a fundamental right, particularly for individuals with disabilities, as recognized by the EU Web Accessibility Directive (WAD)[2]. This directive mandates that public sector websites and mobile applications be accessible, including the requirement for multimedia content to be accompanied by audio descriptions or alternative formats for blind and visually impaired users. However, the practical implementation of these standards across EU member states remains inconsistent and, in many cases, inadequate.
Our presentation discusses results of a study caried out to systematically map the accessibility of multimedia content on the websites and social media platforms of public institutions in Poland and EU bodies.

28 official websites of central government bodies in Poland and the European Union were reviewed to identify embedded multimedia and assess their accessibility for blind individuals. Subsequently, the official social media pages and YouTube channels of these institutions were examined to determine the availability of audio descriptions and alternative access options.

The research highlights a significant gaps and areas for improvement between what is required in theory and what is provided in practice. In Poland, accessible multimedia content is found on the websites of only five out of sixteen ministries, as well as the Central Statistical Office and the State Fund for Rehabilitation of Disabled Persons (PFRON). Notably, the Ministry of Finance and PFRON lead in providing accessible content, with 46 and 47 videos with audio descriptions, respectively, on their YouTube channels. However, this is an exception rather than the norm, and there remains a substantial lack of accessible multimedia across other ministries.

At the EU level, the situation appears even more critical. An analysis of the official YouTube channels of various EU institutions, including the European Commission, European Council, Council of the European Union, and the European Parliament, reveals a complete absence of videos with audio descriptions. Furthermore, the multimedia content on their social media platforms—such as Facebook, Instagram, and X (formerly Twitter)—is similarly inaccessible to blind individuals,

[2] European Union (EU) Directive 2016/2102

lacking alternative text descriptions or audio description that would enable equal access.

This study underscores the urgent need for improved compliance with digital accessibility standards across both national and EU-level public institutions. The absence of accessible multimedia content not only violates the principles enshrined in EU law, but also excludes a significant portion of the population from participating fully in civic life.

Marta Żaczkiewicz is an audio describer, caption writer, and founder of the Open Culture Open Art Foundation, which advocates for individuals with sensory disabilities at risk of social exclusion. A graduate in audiovisual translation studies, she has worked in accessibility since 2010, with a particular interest in the perception of congenitally blind individuals. She has implemented accessibility services for film festivals, theaters, and museums and developed a training program for blind individuals aspiring to be audio description consultants. Since 2023, she has been a PhD student in media accessibility at the University of Antwerp.

Anna Jankowska, PhD, is a Professor at the Department of Translators and Interpreters of University of Antwerp and former Assistant Lecturer in the Chair for Translation Studies and Intercultural Communication at the Jagiellonian University in Krakow (Poland). She was a visiting scholar at the Universitat Autònoma de Barcelona within the Mobility Plus program of the Polish Ministry of Science and Higher Education (2016-2019). Her recent research projects include studies on audio description process, mobile accessibility and software.

A study on the availability and reception of audio description in Iran

Marzieh Izadi
Marzieh.Izadi@student.uantwerpen.be
University of Anwerp

Audio description (AD) in Iran is still in its infancy, both in terms of the provision of the service and as an academic object of study. It is mainly offered in films and thanks to volunteers who write and voice them. Currently, AD service in Iran is only limited to a few Persian and English dubbed movies and there is a big gap between supply and demand. Persian AD practice is facing many difficulties such as deficiency of funding, low quality of AD services, lack of professionals and unified AD guidelines. In order to learn more about the availability of AD and the target audiences' reception of existing AD services in Iran, a study was designed and carried out; the results of which will be presented in this paper. First, the paper provides an overview of the availability of AD in Iran and analyses the difficulties encountered in its development. Then, it conducts a survey among Iranian persons with a visual impairment who have already been exposed to Persian AD. This empirical research shows that the participants are quite satisfied with the quality of AD services provided so far, despite it not being delivered by persons who have been professionally trained in this field. Over interpretation, revealing the plot, subjective speculations, and identifying the source of known sounds are common practices in AD services in Iran but are accepted by the participants since they all help them to understand the film better. This study postulates that the 'novelty factor' (i.e. any AD is better than no access to a film) may partially account for the positive reactions of the Iranian blind and visually impaired audience. The findings of this study offer important insights into the situation of audio description in Iran and recommendations are put forward for future development to serve the Iranian visually impaired community.

Marzieh Izadi is a PhD student in the Department of Applied Linguistics/Translators and Interpreters at the University of Antwerp, Belgium. Her research interests include Audiovisual Translation, Audio Description, and Revision.

The Commercialization of Audio Description in China: Present and Future Prospects through a Case Study of the Yier Company

[1]Dongxue Han, [2]Yuchen Liu
Snow@yier-ea.com
[1]Shanghai Yier Information Co., Ltd
[2]Shanghai International Studies University

China, with the largest population of persons with disabilities (Wu & Xie, 2015), has long relied on volunteer efforts to provide audio description (AD). The most common forms of AD are recorded versions and live sessions, primarily for films (Liu, 2023). However, the absence of legislation mandating AD services has led to significant challenges, particularly related to copyright infringement, which remains a key barrier to the broader development of AD in China (Tor-Carroggio & Casas-Tost, 2020).

In 2021, two critical legal changes created opportunities for the standardization of AD in China: the amended Copyright Law and the ratification of the Marrakesh VIP Treaty. Meanwhile, several AD stakeholders have been actively working on the professionalization and standardization of AD services. Li (2013) had previously proposed a business model to promote accessibility services in China.

This presentation will begin by outlining the current challenges faced by the AD sector in China, particularly concerning its professionalization and standardization. These challenges led to the founding of Shanghai Yier Information Co., Ltd (上海译迩信息技术有限公司) in 2019, the first and only publicly known company in Mainland China to specialize in accessible audiovisual products, including AD, Subtitles for the Deaf and Hard-of-Hearing (SDH), and sign language interpretation. We will provide an overview of the company's development, focusing on its guiding principles and workflows, characterized by a close collaboration between relevant industries, accessibility experts, and end-users. These workflows are tailored to the specific needs of each product, such as those within the film industry, and are applied to a wide variety of audiovisual products. To illustrate these workflows, we will present case studies, including AD services for an augmented reality (AR) tour.

We will conclude by discussing strategies for promoting the AD business model in China, aiming to balance the growing demand for accessibility services with the commercial imperatives necessary for sustainable business operations. Ultimately, this presentation seeks to

provide insights into how AD commercialization could enhance accessibility and inclusion for persons with disabilities in China.

References

Li, D. (李东晓). (2013). *听见,看见:影视媒体的无障碍传播研究 [Hear and See: Audiovisual Media for All].* Hangzhou: Zhejiang Daxue Chubanse.

Liu, Y. (刘禹辰). (2023). Mapping the AD development patterns between China and Europe: China, Spain, and the UK in focus. *Journal of Chinese Film Studies, 3*(2), 241–262.

Tor-Carroggio, I., & Casas-Tost, H. (2020). Who is currently audio describing in China? A study of Chinese audio describer profiles. *MonTI. Monografías de Traducción e Interpretación, 12,* 78–107.

Wu, Z. (吴宗艺), & Xie, Z. (谢桢桢). (2015). 中国视障口述影像服务的发展现状与大陆推广 [The development and promotion of audio description in mainland China]. *Xinwen Yanjiu Daokan, 6*(10), 21.

Founder and CEO of Shanghai Yier Information Co., Ltd, **Dongxue Han** is China's first expert in accessible entertainment and the initiator and leader of the Entertainment Accessibility Project Group, which publicly released the Production Process and Technical Standards of Barrier-free Cultural and Entertainment Works in China for the first time. As the country's first producer of accessible content—including films, advertisements, and e-commerce streaming—she also organized China's first international forum on accessible cinema. Additionally, she initiated and led the country's first professional training program for providing accessible services for people with disabilities. She was an expert speaker for the Accessibility Group at the 72nd annual Berlin International Film Festival. She is also an expert for the China Alliance of Social Value Investment (Shenzhen) and the Information Accessibility Joint Conference (Shenzhen). Furthermore, she has delivered a TEDx talk.

Yuchen Liu is an assistant researcher at the Institute of Translation Studies at Shanghai International Studies University (SISU). She is also a member of the GELEA2LT research group. She holds an international PhD in Translation and Intercultural Studies from the Universitat Autònoma de Barcelona (UAB) (2024) and an M.A. in Translatology and Intercultural Studies from UAB (2019). Her research interests include didactic audiovisual translation for language education, audiovisual translation, and media accessibility.

Audio Description of Spaces and Landscapes for the "SIGA - Accessible Guide of the City" App

Suely Maciel, Guilherme Ferreira de Oliveira, Guilherme Mori Magalhães
suely.maciel@unesp.br
São Paulo State University

This work aims to present the methodology for developing audio descriptions of spaces and landscapes created by the Laboratory of Education, Research, and Extension in Media and Accessibility "Biblioteca Falada" at the São Paulo State University (Unesp), Brazil. Specifically, it highlights its use for the "SIGA - Accessible Guide of the City" app. SIGA is a geolocation app with audio descriptions and historical-informational audios of points of interest in cities in the state of São Paulo, Brazil. The app results from an integrated effort by students, professors, visually impaired consultants, and students and teachers from the Home-School Santa Luzia for the Blind (LESL) in Bauru, Brazil. Audio description (AD) is an intersemiotic translation modality which can be used as an accessibility feature/assistive technology in different settings, such as audiovisual productions, events, publications, theater, etc. Its use is also recognized for enabling access to the visual information of landscapes and spaces. Audio description of spaces and landscapes allows access to the physical characteristics of spaces so that the audience (especially people with visual impairments) can understand them. Urban environments are usually composed of a multiplicity of architectural complexes, buildings, open areas, transit zones, parks, plazas, monuments, and historical sites. Therefore, the AD of spaces and landscapes must consider architectural features, dimensions, shapes, colors, space locations, furniture and object arrangements, and guidelines for movement. The stages for the AD development by Biblioteca Falada are: a) familiarization with the location (through visits, photographs, and videos), b) script preparation, c) script review, d) submission of the script to a visually impaired audio description consultant (an essential step for audio description in the Brazilian context), e) adjustments based on consulting feedback, f) narration, and g) audio editing and production. The scripts are structured in three parts: introduction, contextualization, and development, considering aspects such as style, language, audio description norms, sound text structure, and the creation of mental images of spaces. AD can be provided as audio or written text. In the case of SIGA, the ADs are uploaded on a page for each location in the cities. The information and audio descriptions are available in audio files linked to a geolocation system, allowing users to navigate a map and identify each described place with its location in the cities. Locations in the app can also be accessed via a general list or by categories (such as sports, leisure, culture, religion, etc.). The app

and the ADs it contains are tools for building knowledge about the historical, geographical, social, and cultural points of interest in the cities, contributing to greater access to information and helping people to gain a more autonomous and independent understanding of their lives and their cultural, touristic, leisure, and recreational activities.

Suely Maciel is a professor of undergraduate courses and the Postgraduate Program in Media and Technology at the Faculty of Architecture, Arts, Communication and Design of the São Paulo State University (Unesp). Leader of the Research Group on Language and Accessible Media (Gelima) and Coordinator of the Laboratory of Education, Research and Extension in Media and Accessibility "Biblioteca Falada". Conducts research on the themes of media accessibility, audio description, accessibility features for journalism specialized in tourism, leisure and culture.

Guilherme Ferreira de Oliveira holds a Master's degree in Media and Technology from the Faculty of Architecture, Arts, Communication and Design at São Paulo State University (Unesp). Graduated in Public Relations from the same institution. Member of the Language and Accessible Media Research Group (Gelima) and collaborator of the Laboratory of Education, Research and Extension in Media and Accessibility "Biblioteca Falada".
Conducts research on the topics of media accessibility, audio description, organizational communication and public relations.

Guilherme Mori Magalhães is a master's student in Media and Technology at the Faculty of Architecture, Arts, Communication and Design at São Paulo State University (Unesp). Graduated in Public Relations from the same institution. Member of the Language and Accessible Media Research Group (Gelima) and collaborator of the Laboratory of Education, Research and Extension in Media and Accessibility "Biblioteca Falada". Conducts research on the topics of media accessibility, audio description, video games and digital media.

INCinema: an accessible film festival

[1]Federico Spoletti, [2]Annalisa Sandrelli
federico@subti.com
[1]SUB-TI, SUB-TI Access, FRED Film Radio- The Festival Insider
[2]UNINT - Università degli Studi Internazionali di Roma, SUB-TI

Although in the last few years there has been growing awareness of access issues, film festivals are still largely inaccessible. According to a recent survey on 75 film festivals wolrwide, some of the measures that would improve accessibility include the provision of subtitles for the films and live subtitles for panels and Q&As (ASIR 2023). Despite recent progress, accessible screenings still face challenges, primarily owing to lack of awareness and lack of economic interest among producers, distributors, and exhibitors.

INCinema, launched in Italy in 2023, is a ground-breaking initiative in the realm of film festivals by being the first of its kind to prioritize accessibility for individuals with sensory disabilities. The festival's hybrid format includes both in-person screenings in accessible cinemas across eight Italian cities and online streaming via the MYmovies One VOD platform, thus ensuring broader attendance. Each film is accompanied by Subtitles for the Deaf and Hard of Hearing (SDH) and by Audio Description (AD). Moreover, festival events such as film presentations, debates and so on feature real-time transcription through automatic speech recognition (ASR) or respeaking technology.

The presentation will explore the festival's mission to be more inclusive and allow everyone, regardless of their sensory abilities, to experience the unique magic of film festivals. Furthermore, INCinema also aims to showcase the commercial and social benefits of catering to audiences with sensory impairments. The latter were engaged at various levels: as audience members via user associations (many of which endorsed the festival), via traditional media and via social media platforms. Some people with disabilities were also involved in the festival roundtables and, wherever possible, in reviewing accessible contents. They also provided feedback on the quality of such contents and the overall accessibility of the festival. In addition, collaborations with secondary schools also aim to educate young people about inclusion, fostering a culture of accessibility from a young age.

The presentation will feature case studies of successful international practices, underline the importance of inclusive film festivals, and discuss future directions for making universal accessibility a standard in the film industry. The goal is to inspire stakeholders to recognize the value of inclusive practices and implement them in their own

contexts, thus ensuring that no audience is deprived of the emotional and intellectual enrichment that films provide.

References

ASIR. (2023). *Accessibility Scorecard Impact Report*. A collaborative initiative by Filmmakers with Disabilities (FWD-Doc), the Film Event Accessibility Working Group (FEAW), and the Film Festival Alliance (FFA). Retrieved from https://drive.google.com/file/d/1wpt-bxSQNTZrZjswXYJwML5BJF0PwhP_/view

Born and raised in Italy, **Federico Spoletti** is co-founder and Managing Director of SUB-TI, an international subtitling company based in London, which provides audiovisual translation services all over the world. He is also co-founder of SUB-TI ACCESS, a company involved in cultural accessibility for the hearing and visually impaired, specializing in subtitles for the deaf and audiodescription for the blind. In 2011, Federico launched FRED Film Radio - The Festival Insider. FRED is an internet radio network which boasts 29 channels, broadcasting in 25 languages, with thematic channels dedicated to specific cinema content such as film education and the film industry. In 2023 Federico launched INCinema, the first film festival completely accessible to people with sensory disabilities. He has been involved in the following international projects: FRED at School, LTA, SMART, WICIP and, more recently, FAIR (accessible screenings for the elderly who live in old people's homes).

Prior to joining UNINT (Rome) as a Lecturer in English, **Annalisa Sandrelli** taught in Trieste and Bologna/Forlì and was Marie Curie TMR Fellow and Lector in Italian at the University of Hull. She has published widely on corpus-based interpreting studies, audiovisual translation (dubbing, subtitling, respeaking), EU English, and Computer Assisted Interpreter Training (CAIT). Her most recent publications on AVT include Sandrelli. A. & Bonsignori V. (forthcoming) "Jane Austen's Pride and Prejudice from the Page to the Screen: A Diachronic Analysis of Source Texts and Italian Translations", in I. Ranzato, L. Valleriani(eds) English Classics in Audiovisual Translation, London, Routledge; E. Davitti, A. Sandrelli, T. Korybski, C. Orasan, S. Braun (forthcoming) "Using ASR tools to produce automatic subtitles for TV broadcasting: a cross-linguistic comparative analysis". JOURNAL OF AUDIOVISUAL TRANSLATION; and Sandrelli, A. (2024) "Integrating ASR and MT tools into cloud subtitling workflows: the iSub! and iSub!2 projects", in Y. Peng, H. Huang, D. Li (a cura di) New Advances in Translation Technology. Springer New Frontiers in Translation Studies, Springer Nature Singapore Pte Ltd, Singapore. She has given workshops and seminars on audiovisual translation and accessibility all over Europe, in academic conferences, for film schools and film festivals.

PANEL 6. BUILDING FILMIC EXPERIENCES

Crowdsourcing in Audio Description: The Blue Description Project

María José García Vizcaíno
garciavizcam@montclair.edu
Montclair State University

While the role and impact of crowdsourcing in translation has been thoroughly explored so far (see Jiménez-Crespo, 2022, for a full account on that), research on crowdsourcing in the realm of audio description (AD) is almost non-existent. The aim of this presentation is to share how crowdsourcing has been applied in the Blue Description Project (BDP). In addition, I will examine the challenges of translating this creative audio described script into Spanish for the BDP tour that will take place in several countries in 2025.

The BDP is a collaborative endeavor led by artists and writers Liza Sylvestre and Christopher Jones at University Urbana-Champaign, Illinois (https://faa.illinois.edu/about-us/events/the-blue-description-project/) to create the AD script of the experimental feature film *Blue* (1993). *Blue* is an experimental film directed by Derek Jarman, a well-known British filmmaker and gay-rights activist. In this movie, he reflects upon his battle fighting AIDS against a blue backdrop. For the duration of the entire film (79 minutes), the only visual projected on the screen is a single shot of International Klein Blue. Derek was losing his sight during the same year he was working on this film, only being able to see in shades of blue. This work therefore defies ableist and ocularcentric approaches to filmmaking by using a very visual element such as a color.

The audio descriptions of blue for the BDP use a wide range of linguistic and poetic resources. Some of them are similes either with abstract entities ("A blue that feels like a cold night when you are well-dressed") or concrete elements ("A blue that almost flows on the screen like a river, with purpose"). Others consist of personifications ("A blue that holds your hand, squeezing softly in acknowledgment of you") and metaphors ("A blue of echoes"). And in many occasions, emotions ("A blue that starts to hurt my eyes—I want to look away! Or close my eyes completely!") function as descriptions. This diversity in the AD choices of the same visual input may correlate to the diversity existent in the visually-impaired and blind audiences themselves (Kleege, 2018: 100). Taking this diversity into account, it is only natural that subjectivity has been the leading force in this creation, challenging the norms of standard practice that advocate for neutral and objective AD. In this project, what blind scholar Georgina

Kleege (2022a) claims, more than ever, it is manifested: "Who are you and what are you bringing to the audio description? What does your knowledge bring?" To the listener, that knowledge may be interesting and certainly there is also something empowering in bringing who you are for the collective enjoyment of the group. This type of accessibility is definitely another example of embracing access as collaboration and conversation (Romero-Fresco & Dangerfield 2022), as well as a form of intimate connection (García-Vizcaíno, in press; Mingus, 2011, 2017; Romero Fresco (2024), rather than a mere service to be provided to the blind and low vision community.

References

García-Vizcaíno, M. J. (in press). Love at first touch: A collaborative approach to tactile art explorations. *Journal of Arts and Communities*. Special Issue: "Trans-disciplinarity in Disability, Art, and Design."

Jiménez-Crespo, M. Á. (2022). Quo vadis, crowdsourcing and online collaborative translation? *INContext: Studies in Translation and Interculturalism, 2*(1), 32–56. https://doi.org/10.54754/incontext.v2i1.13

Kleege, G. (2018). *More than meets the eye: What blindness brings to art.* Oxford University Press.

Kleege, G. (2022a). What's it to you?: Subjectivity and audio description. Lecture delivered at Montclair State University, Montclair, New Jersey, 24 February.

Mingus, M. (2011). Access intimacy: The missing link. *Leaving Evidence.* Retrieved from https://leavingevidence.wordpress.com/2011/05/05/access-intimacy-the-missing-link/ (Accessed 6 September 2024).

Mingus, M. (2017). Access intimacy, interdependence and disability justice. *Leaving Evidence.* Retrieved from https://leavingevidence.wordpress.com/2017/04/12/access-intimacy-interdependence-and-disability-justice/ (Accessed 6 September 2024).

Romero Fresco, P., & Dangerfield, K. (2022). Accessibility as a conversation. *Journal of Audiovisual Translation, 5*(2), 15–34.

Romero Fresco, P. (forthcoming). Access intimacy in media accessibility. *Journal of Audiovisual Translation*.

Dr. María José García Vizcaíno is Associate Professor, Director of the Graduate Program in Spanish Translation and Interpretation, and

Chair of the Department of Spanish and Latin Studies at Montclair State University where she teaches courses in translation theory and practice, audiovisual translation, and audiodescription. Her current line of research revolves around audio description, touch, art, and blindness. In particular, she is currently working on a project about tactile resources in art museums, and about the role of touch as an accessibility tool for sculpture. She has published multiple articles in academic journals and has given numerous lectures at national and international conferences.

Experimental audio description as a means of enhancing the cinematic experience of sighted viewers

Iwona Mazur, Julia Kaźmierczak
imazur@amu.edu.pl
Adam Mickiewicz University

Audio description (AD) is an accessibility tool meant primarily for persons with sight loss (PSL), however – according to universalist (e.g., Greco 2016) and functional (e.g., Mazur 2020) approaches to AD – it can also be used by sighted people. While there is plenty of AD reception research involving PSL, similar research with sighted persons is rather scarce (for an overview, see Mazur 2019). In particular, what has not been directly addressed in AD research to date is how sighted viewers respond to experimental AD. The present study is meant to fill in this research gap. It involved sighted participants (n=52) in two experimental groups: the AD and the standard one. The AD group was asked to watch a movie clip with experimental AD (narrated in the first person, cf. Udo and Fels 2009) and fill in a questionnaire concerning (1) the cognitive assessment (cf. Perego 2016), (2) the experiential assessment (the degree of presence; cf. Fryer and Freeman 2013) and (3) the overall evaluation of the quality of the AD. The standard group watched the clip without AD and answered the same questions with the exception of part (3). In the presentation, we will report on results concerning the cognitive and experiential assessment, showing clear benefits of experimental AD for sighted viewers. The study, therefore, adds to the body of research on applications of AD for sighted people, but also opens up a discussion on more creative (non-standard) types of AD, in which the human input may outweigh the possibilities offered by generative artificial intelligence. It will also be suggested, in line with the functionalist approach to AD, how alternative AD styles may ensure not just accessibility, but also greater social inclusion of PSL by making AD more enjoyable and beneficial for sighted persons.

References

Fryer, L., & Freeman, J. (2013). Cinematic language and the description of film: Keeping AD users in the frame. *Perspectives: Studies in Translatology, 21*(3), 1–15.

Greco, G. M. (2016). On accessibility as a human right, with an application to media accessibility. In A. Matamala & P. Orero (Eds.), *Researching audio description: New approaches* (pp. 11–33). London: Palgrave Macmillan.

Mazur, I. (2019). Audio description for all? Enhancing the experience of sighted viewers through visual media access services. In D. Abend-

David (Ed.), *Representing translation: The representation of translation and translators in contemporary media* (pp. 122–149). New York: Bloomsbury.

Mazur, I. (2020). A functional approach to audio description. *Journal of Audiovisual Translation, 3*(2), 226–245. https://doi.org/10.47476/jat.v3i2.2020.139

Perego, E. (2016). Gains and losses of watching audio described films for sighted viewers. *Target, 28*(3), 424–444.

Udo, J. P., & Fels, D. I. (2009). Suit the action to the word, the word to the action: An unconventional approach to describing Shakespeare's *Hamlet*. *Journal of Visual Impairment and Blindness, 103*(3), 178–183.

Iwona Mazur is an Assistant Professor at Adam Mickiewicz University in Poznan, Poland. Her research focuses on audio description, translation theory, and multimodality. She has co-authored a book on AD (Audiodeskrypcja, 2014) and co-edited a special issue of Linguistica Antverpiensia on media accessibility training (2019). She has participated in a number of Polish and international research projects, including ADLAB (2011-2014) and ADLAB PRO (2016-2019). She is the recipient of the 2024 Scientist of the Future Award for her current project, 'Persons with visual impairment and functions of multimodal texts: Development of a functional model of audio description' (NCN, 2023-2026). More information: http://wa.amu.edu.pl/wa/Mazur_Iwona

Julia Kaźmierczak holds a Master of Arts degree in Creative and Specialised Translation from the Faculty of English, Adam Mickiewicz University in Poznan, Poland. Her research interests include audio description, reception research, experimental audio description, and arthouse cinema.

No sound but the AD… How does audio description alter filmic narratives? The case of the Flemish TV series Nachtwacht

Gert Vercauteren
gert.vercauteren@uantwerpen.be
University of Antwerp

In Buhler et al.'s (2010) work on sound and music in fiction film, sound designer Vincent LoBrutto points out that (for sighted viewers) "50 percent of the motion picture experience is aural" (p. xxii). It is safe to say that in the case of the audio described version of this motion picture, that percentage is considerably higher. Yet both in the fields of narratology and audio description, this aural dimension is still rarely taken as the main research focus (Mildorf and Kinzel, 2016; Vercauteren and Reviers, 2022). In the field of AD there have been some small-scale case studies (e.g. Remael, 2012; Vercauteren and Reviers, 2022) and some experimental research studying the narrative potential of sound as an alternative to traditional AD (Lopez and Pauletto, 2009; Lopez, et al., 2020; Lopez, et al., 2021) has been conducted. The results from these endeavours show that sound does have the potential to tell powerful aural stories when combined with AD, but that the audio description also has a considerable impact on the aural dimension of the narrative. Vercauteren & Reviers (2022) found in their case study that less than a third of all sounds were taken into account in the description, meaning that the target audience had to interpret a little over 70% of the sounds without any further explanation.

One fundamental question that so far did not receive any real attention yet, is what impact the addition of an audio description has on the narrative: the AD is added when there is no dialogue, but inevitably will be superimposed on narrative sounds and/or even replace them, which changes the narrative told by the film. In this presentation we want to address this question through an analysis of an episode of the Flemish TV series Nachtwacht (Nightwatch). For this analysis, a new analytical model was developed, building on the initial model developed by Vercauteren and Reviers (2022). A pilot study carried out in the first half of 2024 on two excerpts of the episode, showed that the AD does indeed have an impact on the narrative: in scenes with a lot of dialogue, almost 10% of the sounds were lost because of the AD, in scenes with little dialogue, this increased to 21%.

For the presentation, the analysis carried out in the pilot will be extended to the full episode, and will focus on two main research questions: 1) What is the impact of adding audio description to the aural dimension of the narrative? and 2) To what extent does the audio description take sound into account cf. earlier research by

Vercauteren and Reviers). On the one hand, the findings will contribute to existing knowledge in the field of audionarratology, on the other they can have both theoretical and practical implications for the field of audio description.

References

Buhler, J., Neumeyer, D., & Deemer, R. (2010). *Hearing the movies: Music and sound in film history*. Oxford University Press.

López, M., Kearney, G., & Hofstädter, K. (2020). Seeing films through sound: Sound design, spatial audio, and accessibility for visually impaired audiences. *British Journal of Visual Impairment*.

López, M., Kearney, G., & Hofstädter, K. (2021). Enhancing audio description: Inclusive cinematic experiences through sound design. *Journal of Audiovisual Translation, 4*(1), 157–182.

López, M., & Pauletto, S. (2009). The design of an audio film: Portraying story, action, and interaction through sound. *JMM: The Journal of Music and Meaning, 8*, Winter 2009.

Mildorf, J., & Kinzel, T. (2016). Audionarratology: Prolegomena to a research paradigm exploring sound and narrative. In J. Mildorf & T. Kinzel (Eds.), *Audionarratology: Interfaces of sound and narrative* (pp. 1–26). De Gruyter.

Remael, A. (2012). For the use of sound: Film sound analysis for audio description. Some key issues. *Monti: Monografías de traducción e interpretación, 4*, 255–276.

Vercauteren, G., & Reviers, N. (2022). Audio describing sound: What sounds are described and how? Results from a Flemish case study. *Journal of Audiovisual Translation, 5*(2), 114–133.

Gert Vercauteren is a tenure track lecturer at the University of Antwerp and a member of the TricS Research Group and the OPEN Expertise Centre for Accessible Media and Culture. His holds a PhD in translation studies, focused on content selection in audio description. His current research interests include the role of sound in audio description and narratology more in general, and the impact of audio description on cognitive load. In addition to his position at the university, he works as an audio describer for film and television in Flanders.

Creativity in audio description of the horror genre: a reception research

Marcella Wiffler Stefanini
marcella.wiffler@gmail.com
University of Campinas

Audio description (AD) is the translation of (audio)visual texts into verbal texts with the aim of making them accessible to people with disabilities (Fryer, 2016). Considering it, therefore, as an intersemiotic translation, the research seeks to study the possibilities of audio describing three horror scenes from the short film "Enjaulado" (1997), by Kleber Mendonça Filho. In previous research (Stefanini & Matamala, 2023), we could verify there was a lack of studies on the AD of horror films, which poses several challenges to the audio describer, especially because horror films use many visual elements to frighten and scare the viewer (Cherry, 2009), and also because cinema commonly combines visual and audio information, which, in the case of horror films, complement each other in the construction of a horror atmosphere (Santos, 2018). Such particularities make the audio describer's work more challenging, as the AD must avoid overlapping the sound content considered relevant to the narrative, such as dialogues and other diegetic sounds, and therefore needs to be concise and inserted before or after this information (Franco & Araújo, 2011). In addition to deciding "what" and "when" to audio describe (whether before or after the sound elements), an important decision that the audio describer must make, and perhaps the most complex, is "how" to translate. Given these challenges, we proposed in this research two different ADs for the same scenes, with the aim of investigating whether translation solutions considered more creative (considering specially the "how") and whether the moment of insertion of the AD (before or after the sound effect) have an impact on the reception of horror scenes by the visually impaired audience. To this end, we developed two ADs versions for the selected scenes of the short film: one based on the guidelines of the manuals (we used as main references the Brazilian manuals "Guia para Produções Audiovisuais Acessíveis" and "ABNT NBR 16452", both from 2016), and another that we called the creative version, more concerned with translating the (audio)visual language responsible for creating the horror atmosphere. The main difference between both approaches is that, due to the lack of dialogues – once the film is centered in one character –, there was enough time to provide a more detailed description, which we have done in the first AD. Nonetheless, most of the film's tension is built from the sound effects, which made us experiment with shorter descriptions in order to preserve the audio information for the "creative" version. Another difference between the two lies in the AD's narration. Although the "non-creative" version has a vivid

narration, the “creative” one sought a narration that would complement the sound effects, through voice modulation and its intensity. A reception research (Chmiel & Mazur, 2012) will now be conducted in order to verify the preference of visually impaired audiences between the two AD versions, and hopefully presented in ARSAD 2025.

Marcella Wiffler Stefanini is a PhD student in the Graduate Program in Applied Linguistics (PPG-LA) at the University of Campinas (Unicamp), in the research line Languages, Transculturality and Translation. She has a master's degree in Applied Linguistics (2020) and a degree in Letters – Portuguese Language (2017), also from Unicamp. She is focused on studying Audio Description as an Intersemiotic Translation since her master’s thesis, and on her PhD dissertation studies the possibilities of audio describing horror films. As a PPG-LA student, she participates in the research group E por falar em Tradução, which encompasses research on Translation Studies in its various spheres. Her other research interests include Audiovisual Translation, Audio Description and Cinema Studies.

AD PITCHES II

Professional vs Trainee Translators and Audiodescription: an exploratory study on taboo images audiodescribed into Portuguese

[1]Catarina Xavier, [2]Cláudia Martins
catarina.andradexavier@gmail.com
[1]Faculdade de Letras da Universidade de Lisboa
[2]Instituto Politécnico de Bragança

This study reports on a case study on the differences and similarities in audiodescription produced by professional and trainee translators when addressing audiovisual products with taboo subjects, such as sex, violence and religious images in audiovisual content. There is limited research on the accessibility of taboo(s), especially in audiodescription, with most existing studies being either product-oriented (Franco and Monteiro 2013; Villela 2017; Sanz-Moreno 2017, 2018a, 2018b) or reception-oriented (Sanz-Moreno 2020; Ramos Caro, Espín López and Rojo López 2021; Rojo López, Ramos Caro and Espín López 2022) that are worth highlighting. This study aims to fill this gap in the field of Audiovisual Translation and Media Accessibility and gain insight into how professionals and trainees deal with taboo images in their work, as they are increasingly common in fictional media. Accordingly, we propose exploratory research to address this gap by investigating the different practices of professionals and trainees regarding taboo accessibility for audiences with visual disability in Portugal. The study will be conducted using corpus analysis to examine the treatment of taboos in audiodescription by professional translators working for the streaming giant Netflix, and by trainee translators in two different universities in Portugal. Furthermore, this study also aims to understand if the type and subtype of sex, violence and religious images are relevant variables for the audiodescription of taboos by professional and trainee translators.

Catarina Xavier holds a PhD in Translation Studies and her current research looks into taboo accessibility in audio description and subtitling for the Deaf and hard of hearing. By focusing on the fact that sensory-impaired users are entitled to be given access to audiovisual products and meanings, she is currently examining users' needs and actors' attitudes to impact on academia and society. She has been awarded doctoral and post-doctoral grants, and in the meantime has contributed to the area of the translation of taboo in different media with several academic outputs, promoting the study of translation norms and the systematisation of translation strategies.

Cláudia Martins holds a PhD in Translation by the University of Aveiro, with a thesis on museum accessibility for people with visual impairment. She was awarded a Diploma of Advanced Studies in Translation and Intercultural Studies (Spain) and master in Terminology and Translation and bachelor's degree in Modern Languages and Literatures, both at the Faculty of Arts and Humanities of the University of Porto, Portugal. Since 2001, she has been teaching English as a foreign language, English Linguistics, Terminology and Audiovisual Translation at the School of Education in the Polytechnic Institute of Bragança, Portugal. Her academic interests include Terminology, Phraseology and Paremiology, Audiovisual Translation, with a special focus on Media Accessibility.

What (more) can we do with words: An interdisciplinary workshop on creating Museum Audio Descriptions with the use of Generative AI

[1]Theodora Psoma, [2]Thodora Sazo, [2]Christina Fragkou, [2]Kleoniki Nerantzoglou
tpsom@enl.auth.gr | mea23023@uom.edu.gr | mea23005@uom.edu.gr | mea23023@uom.edu.gr
[1]Aristotle Univeristy of Thessaloniki (AUTh)
[2]University of Macedonia

Accessibility practitioners use words as raw materials when creating Audio Description texts (AD). This is the main focus of our study, the power of words in creating accessible and inclusive content for museums.

Museums, as Graham Black argues, are expected to be "agents of physical, economic, cultural and social regeneration, accessible to all" (2005). In the case of Blind and Partially Blind (BPB) audiences, the need for accessibility leads to the inclusion of Audio Description as an integral part of museums (Perego & Taylor, 2022). Despite this fact, there is scarce progress in creating a framework for producing AD texts. In this context, our workshop aspires to promote the use of writing skills together with Generative AI in order to provide sustainable strategies regarding AD training.

Our presentation for ARSAD 2025 discusses the outcomes and further possibilities of our interdisciplinary AD workshop, delivered on ICOM International Museum Day 2024 at the Jewish Museum of Thessaloniki. As a research group our aim was twofold; on one hand to explore and assess the museum's accessibility strategies and infrastructure and on the other hand to provide training for prospective Audio Describers in creating engaging AD texts with the use of AI. The workshop took place under the ICOM 2024 theme "Museums, Education, and Research" with the participation of 12 undergraduate students coming from various academic fields: Translation Studies, Linguistics, Theatre & Film Studies, Education, Disability Studies, Psychology.

The AD workshop was divided in two parts: Firstly, we delivered a theoretical training on assessment strategies for measuring accessibility for BPB audiences. We focused on the role of Assistive Technology, tactile aids and AD (Vaz et al., 2022; Avni et al., 2024). In the second part we proceeded with the hands-on practice on AD creation by selecting exhibits from the museum's permanent collection and writing texts in groups. Participants' training needs and attitudes towards the workshop methodology were assessed through a standardized questionnaire.

Our results showcase that there is indeed a need for a sustainable and interdisciplinary framework on which Audio Describers can be trained, while the use of AI can be of important assistance.

References

Avni, Y., Danial-Saad, A., Sheidin, J., & Kuflik, T. (2024). Enhancing museum accessibility for blind and low vision visitors through interactive tangible user interfaces. *SSRN*. https://doi.org/10.2139/ssrn.4824710

Black, G. (2005). The engaging museum. In *Developing museums for visitor involvement*. London-New York: Routledge.

Perego, E., & Taylor, C. (2022). A profile of audio description end-users: Linguistic needs and inclusivity. In *The Routledge Handbook of Audio Description* (pp. 38-54). Routledge. Accessed on: 13 April 2023. https://www.routledgehandbooks.com/doi/10.4324/9781003003052-5

Vaz, R., Freitas, D., & Coelho, A. (2022). Enhancing the blind and partially sighted visitors' experience in museums through integrating assistive technologies, multisensory and interactive approaches. In *Lecture Notes in Computer Science* (pp. 521–540). https://doi.org/10.1007/978-3-031-05039-8_38

Theodora Psoma – PhD candidate, Aristotle University of Thessaloniki.
I am a PhD candidate in Applied Linguistics at the School of English, Aristotle University of Thessaloniki. The focus of my research is the language of Museum Audio Descriptions. In my thesis I explore linguistic aspects of accessibility, collaborating with Blind and Partially Blind audiences, as well as with people with various learning (dis)abilities.
For more than 15 years I have been studying EFL while teaching writing skills to students of all ages and all levels of education. As an educator, writer and translator, I have been a member of many creative groups, collaborating with artists in the process of curating texts for exhibitions.
My latest creative project combines the prose of Sylvia Plath with Performative Arts and it is under publication by Routledge.
Research interests: EFL, Corpus Linguistics, Accessibility Studies, Post-Colonial Theory, Translation Studies, Creative writing.

Thodora Sazo – I am a Special Education Teacher with a diverse academic background. I hold a degree in Film Studies from Aristotle University of Thessaloniki and I am currently pursuing an MSc in

Educational Sciences: Special Education and Rehabilitation at the University of Macedonia. My specialization is in the education of individuals with visual impairment. During this period, I am actively engaged in research on an Accessibility project at the University of Macedonia aimed at improving access for people with disabilities. As both a dedicated special education teacher and a researcher, my interest focuses on promoting inclusive education and ensuring accessibility across various sectors, particularly in education and cultural heritage. I firmly believe in lifelong learning and I continuously seek to expand my knowledge through seminars, workshops, and professional development opportunities.

Christina Fragkou – Master of Science in Special Education and Rehabilitation. I am a professional in Physical Education and Sports Science, specialized in Special-Adapted Physical Education. Currently, I am attending a Master of Science in Special Education and Rehabilitation, with a focus on Visual Impairment. My passion lies in advancing inclusive educational practices, particularly on disabilities. As an Adapted Physical Education teacher and a Lead Trainer at the Swiss International Leadership Camp & Academy, I have the opportunity to lead school and summer camps around the world, focusing on creating accessible and engaging physical activities. I am dedicated to continuous professional development, regularly attending workshops on disabilities, sports case studies, and new technologies. My interests include the development and application of tactile models and assistive technologies to support visually impaired individuals.

Kleoniki Nerantzoglou – I am a Kindergarten Teacher with a passion for promoting inclusion in education. I am currently pursuing a Master's degree in Special Education and Rehabilitation, specializing in Visual Impairment. My lifelong interest for crafts and creativity led me to join a university project called Open Book, which involves creating originals or transforming classic children's literature texts into tactile illustrated books. The project was awarded in 2023 by the Greek section of the International Board on Books for Young People, for promoting literacy. This sparked my growing interest in tactile representation and accessibility, encouraging me to explore novel ways to make both education and everyday life of visually impaired individuals more accessible and inclusive. My aim is to become more actively involved in the field of visual disability, tactile representation and accessibility in general.

Audio Description in the Era of AI: Challenges and Future Development

Jackie Xiu Yan, Su Lin
ctjackie@cityu.edu.hk
City University of Hong Kong

Audio Description (AD) provides accessibility service to people with visual impairment. AD has received increasing attention from both the public and academia over the past 20 years, as a result, research and practice in this field have developed rapidly. Given the high degree of similarity between the processes of translating between languages and translating visual content into words, it is possible to apply the principles of translation studies (TS) to AD, consequently, AD has been considered a form a translation (audio-visual translation or intersemiotic translation). Just like the industry of translation, AD is facing the challenges brought by the recent development in AI. For example, technological innovations such as ChatGPT, automatic speech recognition, machine translation have dramatically changed the prospect of translation and AD studies. This study, based on a review of AD-related technological development and an experiment comparing AD output generated by human professionals and AI-powered tools, attempts to examine the challenges and possible direction for future development in this field. A sample of video slips is chosen for the production of both human-created and AI-generated AD transcripts, which will be evaluated by a group of 4 researchers consisting of both sighted and visually impaired people. Factors such as choice of details, language use, contextual appropriateness, and overall user satisfaction will be discussed. The results will provide insights into the strengths and limitations of human-generated and AI-based AD, and will shed light on how AI-powered AD may complement or augment human efforts, focusing on opportunities for collaboration and workflow optimization.

Funding information

The work described in this paper was partially supported by the Hong Kong RGC Grant CityU 11609621 and DON-RMG Grant9229125.

Prof Jackie Xiu Yan received her PhD from the University of Texas at Austin, USA. She is now Subject Leader of the MA Translation Program in the Department of Linguistics and Translation at City University of Hong Kong. Her research and teaching interests include interpreter and translator training, audio description, Chinese culture and translation, and Applied Linguistics. She is now President of the Hong Kong Audio Description Studies Association (HKADSA), President of the Hong Kong Association of University Women

(HKAUW). In 2021, she has won a silver medal in the International Exhibition of Inventions of Geneva.

Ms. LIN Su is a PhD student at the Department of Linguistics and Translation, City University of Hong Kong. Her research interests lie in the areas of public policy, audio description, translation studies, and applied linguistics.

Characters, Scenarios and Objects Audio Description in Games: a study with Mortal Kombat 1

Guilherme Mori Magalhães, Suely Maciel
guilherme.m.magalhaes@unesp.br
São Paulo State University (Unesp), Brazil

Audio description is an important tool when it comes to providing access to cultural and audiovisual environments. Among the different scenarios where audio description can be beneficial to users are video games. Recognized nowadays beyond their therapeutic and educational roles, games also offer opportunities for socialization, subjective and cultural expression for their players, both for the non-disabled community and for players with disabilities (Wästerfors & Hansson, 2017). Although other accessibility features are already in full use in games, such as subtitling, dialogue transcripts, and customization of controls and input methods, research on the potential of implementing audio description as a means of access to visual and aesthetic information is still in its early stages (Larreina-Morales & Mangiron, 2023), with challenges and potentials being pointed out by academia. Cutscenes, trailers, demos, HUDs and GUIs are some of the contexts where audio description can be implemented and its benefits can be enjoyed by a broad audience (Mangiron & Zhang, 2022). Looking at the academic research and titles released on the market in recent years, it is possible to find examples that show the exploration of these audio description possibilities, such as the audio description of cutscenes in The Last of Us Part I and Mortal Kombat 1, the trailer of Assassin's Creed Valhalla, and several other independent games developed for research purposes. By analyzing such examples, this study aims to expand the ways in which audio description is being experimented with on video games, proposing a model for describing characters, scenarios, and objects to be applied as an audio introduction on the game menu. We believe that, in this way, we can contribute to the apprehension of aesthetic information about these games, enhancing the experience of immersion and gameplay. On a methodological standpoint, a design thinking process is currently underway, coming from a bibliographic and documental research aligned with theoretical and reception studies, and the participation of disabled players. Adapting an existing AD model used for describing soap opera characters that was already used in Brazil (Maciel & Caruso, 2017), we pinpointed some general guidelines for this audio introduction, considering also the usage of the characters' original voice and the game's soundtrack, in order to optimize the listener's ability to recall the information during gameplay. A first prototype of audio introduction script and means for its implementation was developed using the game Mortal Kombat 1 as a base. Our next step is conducting focus groups (Matamala, 2021) with disabled gamers in

order to gather inputs and feedback on the content and length of the audio introduction. We intend to proceed with a further prototype iteration to reach a model that could be experimented on a larger scale with different game titles and genres.

References

Larreina-Morales, M. E., & Mangiron, C. (2023). Audio description in video games? Persons with visual disabilities weigh in. *Universal Access in the Information Society*.

Maciel, S., & Caruso, F. N. O. (2017). Audiodescrição de personagens como recurso de acessibilidade à telenovela brasileira. In L. M. Villela (Ed.), *Acessibilidade audiovisual: Produção inclusiva nos contextos acadêmicos, culturais e nas plataformas web* (pp. 113-126). Canal6 Editora.

Mangiron, C., & Zhang, X. (2022). Video games and audio description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 376-389). Routledge.

Matamala, A. (2021). Qualitative research methods in Media Accessibility: Focus groups and interviews. *LEAD ME Summer Training School*.

Wästerfors, D., & Hansson, K. (2017). Taking ownership of gaming and disability. *Journal of Youth Studies*, 20(9).

Guilherme Mori Magalhães – Master's student in Media and Technology at the Faculty of Architecture, Arts, Communication and Design at São Paulo State University (Unesp). Graduated in Public Relations from the same institution. Member of the Language and Accessible Media Research Group (Gelima). Conducts research on the topics of media accessibility, audio description, video games and digital media.

Suely Maciel – Professor of undergraduate courses and the Postgraduate Program in Media and Technology at the Faculty of Architecture, Arts, Communication and Design of the São Paulo State University (Unesp). Leader of the Research Group on Language and Accessible Media (Gelima) and Coordinator of the Laboratory of Education, Research and Extension in Media and Accessibility "Biblioteca Falada". Conducts research on the themes of media accessibility, audio description, accessibility features for journalism specialized in tourism, leisure and culture.

PANEL 7. MUSEUM AND ART AUDIO DESCRIPTION

Accessibility for All in Museums: Applying Technology to Audio Description

Martina Maggi
martina.maggi@unito.it
University of Turin; University of Genoa

In 1999, the then-director of the Whitney Museum of American Art in New York Maxwell L. Anderson wondered about the possible consequences of introducing technology in museums, an environment that had always been patient, tangible, and relaxed.
Indeed, with the advent of the World Wide Web, museums did lose their authority and exclusivity in favour of a democratisation that changed the role of art in society (Besser 2019). Art was then accessible to an increasingly larger amount of visitors from all around the world.

Consequently, the approachability offered by technology raised a new and growing demand for accessibility for all. The request was reinforced by the arrival of a technological barrier alongside the pre-existing physical one (Lisney et al., 2013). In fact, most technological solutions did not provide cognitive, physical, or sensory access (Gilbert 2016; Freeman et al. 2016; Vaz, Fernandes, and Veiga 2018).

In the last thirty years, targeted institutions and organisations, as well as Law Acts, tirelessly addressed the topic of accessibility and inclusivity in the arts. Accessible architecture, accessible websites, and exhibitions flourished.

For blind and visually impaired visitors, who are the focus of the present study, audio description tracks, QR codes, immersive experiences, and devices were studied, designed, and created.

However, even a QR code could represent a barrier between the visitor and the audio described work of art, as the code needs to be detected and scanned with the mobile camera focused.

Two British museums, the Science Museum in London and the Pitts River Museum in Oxford, proposed different approaches.
Respectively, a tracking system developed with iOS technology and incorporated in the hand-held devices, and the NaviLens QR Code, which can be scanned from a longer distance and does not need a hold focus of the camera.

While the Science Museum incorporated the tracking system in the audio description tracks, the NaviLens QR codes formed part of an audio-guided tour named "Nothing Without Us: Experiences of Disability".

The final aim of the presentation is to analyse the proposed approaches and address the potential benefits of employing technology to expand inclusivity in the arts. In particular, it focuses on the development of new technological features designed for blind and visually impaired people and their potential application to permanent audio description tours, with an eye toward accessibility for all.

References

Besser, H. (2019). The future of museums: How will they evolve due to digital changes and in relation to time-based media. In O. Grau, J. Hoth, & E. Wandl-Vogt (Eds.), *Digital art through the looking glass: New strategies for archiving, collecting, and preserving in digital humanities* (pp. 247–260). Edition Donau-Universität.

Freeman, A., Adams Becker, S., Cummins, M., McKelroy, E., Giesinger, C., & Yuhnke, B. (2016). *NMC Horizon Report: 2016 Museum Edition*. New Media Consortium. https://doi.org/10.1111/j.2151-6952.2010.00064.x

Gilbert, S. (2016). Please turn on your phone in the museum: Cultural institutions learn to love selfies and social media. *The Atlantic*, 3, 32–33. https://www.theatlantic.com/magazine/archive/2016/10/please-turn-on-your-phone-in-the-museum/497525

Lisney, E., Bowen, J. P., Hearn, K., & Zedda, M. (2013). Museums and technology: Being inclusive helps accessibility for all. *Curator - The Museum Journal*, 56(3), 353–361. https://doi.org/10.1111/cura.12034

Vaz, R., Fernandes, P. O., & Veiga, A. (2018). Interactive technologies in museums: How digital installations and media are enhancing the visitors' experience. In J. Rodrigues, C. Ramos, P. Cardoso, & C. Henriques (Eds.), *Handbook of research on technological developments for cultural heritage and eTourism applications* (pp. 30–53). IGI Global. https://doi.org/10.4018/978-1-5225-2927-9.ch002

Martina Maggi is a translator, interpreter, and PhD candidate at the University of Turin, Italy. She holds a Bachelor's Degree in Theories and Techniques of Interlingual Mediation (2017) and a Master's

Degree in Translation and Interpreting (2020) from the University of Genoa, as well as an MA in Audiovisual Translation in Cinema, Television, and Multimedia Publishing (2022) from the University of Turin. She is in her third and final year of PhD in Digital Humanities, focusing on audio description and accessibility for all in the arts and museums. Between 2023 and 2024, she undertook her research abroad, in London and Oxford, for a six-month internship with VocalEyes. Her research interests include audiovisual translation, audio description, and translation theory.

A multilingual exploratory study of generative AI usage for museum audio description

[1]María Isabel Rivas Ginel, [2]Martin Kappus, [2]Alexa Lintner, [1]Joss Moorkens, [3]Will Noonan, [1]Lucía Pintado-Gutiérrez
isabel.rivasginel@dcu.ie
[1]Dublin City University
[2]Zurich University of Applied Sciences
[3]Université de Bourgogne

In recent decades, thanks to growing initiatives towards more inclusive and accessible spaces, audio description (AD) has flourished until finally becoming a field of study on its own and a commercial venue for the private sector (Taylor & Perego, 2022). Furthermore, where once research was mostly focused on film and television, the number of papers analysing the characteristics that set museum AD apart has also increased (cf. e.g. Fresno 2022). To address time-consuming and cost-intensive AD production processes, scholars have focused on AD software (ADLAB PRO, 2017; Mendoza & Matamala, 2019; Minutella, 2022) and automation to increase the percentage of accessible content (Wang et al., 2021; Yousif & Al-Jammas, 2023; Doore et al., 2024).

With the advent of generative AI (genAI) technologies and the democratisation of large language models (LLMs), it becomes crucial to study their potential usefulness in the field of museum AD given the medium's fluidity and less contained nature (Braun & Starr, 2022). Therefore, this communication will present the results of an exploratory study comparing image descriptions created by Anthropic's LLM Claude with those from students initiated in audio-describing practices and those from professional audio describers, which will be set as the golden standard. This international endeavour involves students from three different institutions who will work towards their native languages (French, German, Spanish, and English). By contrasting Claude's output with that of novices and experts in a multilingual context, this project aims to shed light on the potential of genAI in museum AD as well as to pinpoint essential differences between the three samples to set a base for future research.

References

ADLAB PRO. (2017). *Assessment of current AD training practices*. Project report no. 1. Retrieved from http://www.adlabproject.eu

Braun, S., & Starr, K. (2022). Automating audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio*

description (pp. 391–406). Routledge. https://doi.org/10.4324/191371

Doore, S. A., Istrati, D., Xu, C., Qiu, Y., Sarrazin, A., & Giudice, N. A. (2024). Images, words, and imagination: Accessible descriptions to support blind and low vision art exploration and engagement. *Journal of Imaging*, 10(1), 26. https://doi.org/10.3390/jimaging10010026

Fresno, N. (2022). Research in audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 312–327). Routledge. https://doi.org/10.4324/191371

Mendoza, N., & Matamala, A. (2019). Skills and competences of audio describers in Spain. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 144–165.

Minutella, V. (2022). Audio description software tools. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 331–352). Routledge. https://doi.org/10.4324/191371

Taylor, C., & Perego, E. (Eds.). (2022). *The Routledge handbook of audio description*. Routledge.

Wang, Y., Liang, W., Huang, H., Zhang, Y., Li, D., & Yu, L. F. (2021). Toward automatic audio description generation for accessible videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). ACM. https://doi.org/10.1145/3411764.3445290

Yousif, A. J., & Al-Jammas, M. H. (2023). Exploring deep learning approaches for video captioning: A comprehensive review. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 100372. https://doi.org/10.1016/j.eprime.2023.100372

María Isabel Rivas Ginel is a postdoctoral researcher at the School of Applied Language and Intercultural Studies and the ADAPT Centre in Dublin City University (DCU). Her current position involves working on the topics of translation technology, large language models and generative artificial intelligence (genAI), translators' attitudes towards new technologies, and accessibility. She holds a PhD in language science delivered by the Université de Bourgogne France-Comté (Dijon, France) and her fields of interest are audiovisual and multimedia translation, translation technology, gender studies, accessibility, and inclusivity.

Martin Kappus has been a lecturer at the ZHAW School of Translation and Interpreting for 15 years. He holds a PhD from Stony Brook University. Before joining ZHAW, he gained valuable

experience in the translation industry as a consultant for a CAT-Tool manufacturer and a large language service provider. His research and teaching interests focus on language technology, with particular emphasis on translation technology and accessibility studies. He also provides live audiodescriptions for soccer matches in Switzerland.

Alexa Lintner is a Research Associate in Accessibility Studies at the Zurich University of Applied Sciences, where she completed her MA in Applied Linguistics in 2018. Her Master's thesis was awarded the 2018 Lionbridge Prize. Alexa's research interests focus on audio description and text simplification. She is currently working on the large-scale, Innosuisse-funded research project entitled "Inclusive Information and Communication Technologies" (IICT) aiming to radically change production workflows in accessible communication, i.e. from human-based to AI-assisted to fully automated. Alexa co-founded the Zurich-based traduko cooperative, where she works as a part-time translator, audio describer and subtitler.

Joss Moorkens is an Associate Professor at the School of Applied Language and Intercultural Studies in Dublin City University (DCU), Science Lead at the ADAPT Centre, and member of DCU's Institute of Ethics and Centre for Translation and Textual Studies. He has published over 60 articles and papers on the topics of translation technology interaction and evaluation, translator precarity, and translation ethics. He is General Coeditor of the journal Translation Spaces, coeditor of a number of books and journal special issues, and coauthor of the textbooks Translation Tools and Technologies (Routledge 2023) and Automating Translation (Routledge 2025). He sits on the board of the European Masters in Translation Network.

Will Noonan is a senior lecturer in English and translation studies at the University of Burgundy in Dijon, France, where he coordinates MA programs in multimedia translation and media accessibility. He is a member of the Centre Interlangues - Texte Image Langage research centre at his home university and of the translation technologies and audiovisual translation working groups within the EMT network. Originally a specialist in humour studies, his research increasingly focuses on audiovisual and digital localisation and media accessibility.

Lucía Pintado Gutiérrez is a lecturer at the School of Applied Language and Intercultural Studies in Dublin City University. She holds a European PhD on translation in language education from the University of Valladolid (Spain) where she completed a BA in Translation and Interpreting and an MPhil in Translation and Intercultural Communication. Her areas of research interest and expertise include media accessibility; audiovisual translation in language education; sociolinguistics and translation; translation, memory and conflict.

Large Language Models as a Potential Source for Automatic Audio Description of Art in Polish. The Case of the Be My AI Tool powered by GPT-4 and GPT-4o

Weronika Magdalena Kostyra
w.kostyra11@gmail.com
The Institute of Applied Linguistics (ILS), University of Warsaw

In spite of being commonly viewed as a factor which limits human imagination and creativity, large language models (LLMs) can turn out to be an important source of image descriptions. They can even become a sort of image provider due to their ability to recreate the uploaded images through words (Chmiel & Mazur 2014). Large language models seem to be especially beneficial for the blind or low sighted people, enabling them to actively perceive the world by means of generated descriptions of the surrounding reality (Perego 2019). The aim of this presentation is to ascertain whether the implementation of the new version of GPT-4o has improved LLM tools' efficiency, accuracy and quality in the machine audio description (AD) of paintings in Polish. The comparison will be based on the analysis carried out using Be My AI tool powered by GPT-4o and its previous GPT-4 version. The objective of the original study (Kostyra 2024) covering the GPT-4 iteration was to ascertain the quality of descriptions of paintings belonging to different artistic movements in both: Polish and Spanish, according to the best AD practices described in the subject literature (Fundacja Kultury bez Barier 2012; NIMOZ 2015; Szarkowska 2009; AENOR 2005; ADLAB 2012; Giansante 2003; Castro Fernández et al. 2018; Künstler, Butkiewicz, Więckowski 2012). In this presentation, five paintings out of original collection of 15 works of art will be analysed. These were selected from the sample audio described using GPT-4 algorithm (Kostyra 2024), thus allowing for a comparison between the two models. The conclusions of the current analysis will allow to determine whether and to what extent the GPT-4o version proves to be more efficient and accurate than its predecessor. Moreover, the final part of the presentation will reveal to what extent the GPT-4o differs from its previous iteration. It will also demonstrate whether the application of the GPT-4o indicates new needs or shortcomings that should be considered while conducting further studies. The topic of this presentation is extremely important as it gives a critical look at emerging new technologies, especially in the field of accessibility. There is a growing interest among the community of blind and low sighted people with regard to the increasing use of LLMs and their continuous improvement (Jankowska 2008). Furthermore, it is hoped that such constant development will make members of this community more independent and actively involved in numerous areas of life, including in the sphere of art.

References

ADLAB: Audio Description: Lifelong Access for the Blind. (2012). *Report on user needs assessment*. Retrieved from http://www.adlabproject.eu/Docs/WP1%20Report%20DEF

AENOR. (2005). *Norma UNE 153020: Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías.* Madrid, Spain: AENOR.

Castro Fernández, D., Cosío Piccone, G., Hu Huang, D., & Calderón Díaz, S. (2018). Evaluación del proceso y producto de una propuesta de audiodescripción de tres obras de la exhibición permanente de arte moderno del museo de arte de Lima. In *SENDEBAR. Artículos originales* (No. 29, pp. 147–177). Universidad Peruana de Ciencias Aplicadas. ISSN-e 2340-2415.

Chmiel, A., & Mazur, I. (2014). *Audiodeskrypcja*. Poznań, Poland: The Faculty of English of Adam Mickiewicz University.
Fundacja Kultury bez Barier. (2012). *Audiodeskrypcja – zasady tworzenia*. Warsaw, Poland. Retrieved from https://kulturabezbarier.org/wp-content/uploads/2019/12/Audiodeskrypcja-zasady-tworzenia.pdf

Giansante, L. (2003). *Writing verbal descriptions for audio tours.* Art Beyond Sight. Retrieved from https://moodle2.units.it/pluginfile.php/556128/mod_resource/content/3/Giansante%202015%20Writing-for-Audio-Guides-short.pdf

Jankowska, A. (Coord.). (2008). *Między oryginałem a przekładem. Wzniosłość i styl wysoki w przekładzie* (Vol. XIV). Cracow, Poland: The Jagiellonian University.

Kostyra, W. M. (2024). Los grandes modelos de lenguaje como potencial fuente de la audiodescripción automática de arte: El caso de las herramientas Copilot y Be My Eyes/AI en las versiones polaca y española. Warsaw, Poland: The Institute of Applied Linguistics, University of Warsaw.

Künstler, I., Butkiewicz, U., & Więckowski, R. (Eds.). (2012). *Audiodeskrypcja – zasady tworzenia*.

Narodowy Instytut Muzealnictwa i Ochrony Zabytków (NIMOZ). (2015). *Rekomendacje dotyczące udostępniania instytucji muzealnych osobom z niepełnosprawnością wzroku i tworzenia audiodeskrypcji do dzieł plastycznych* (Ed. B. Szymańska). Poland: Narodowy Instytut Muzealnictwa i Ochrony Zabytków.

Perego, E. (2019). Gains and losses of watching audio described films for sighted viewers. *Archivo della ricerca*. Università degli Studi di Trieste, Italy.

Szarkowska, A. (2009). Audiodeskrypcja oczami niewidomych – wywiad z Tomaszem Strzymińskim. In *Przekładaniec* (No. 20, pp. 136–137). Cracow, Poland.

Weronika Magdalena Kostyra – Graduated with distinction at the Institute of Applied Linguistics, University of Warsaw. Student of the School of Spanish Law at the Faculty of Law and Administration, University of Warsaw; Associate Academic Teacher at the Institute of Applied Linguistics, University of Warsaw; Member of NKPL (Literary Translation Research Group) at the Faculty of Applied Linguistics, University of Warsaw. Her MA thesis was devoted to the evaluation of large language models' capacity to generate machine audio description. Her diverse academic interests include modern technologies in audio description and literary translation.

PANEL 8. DANCE AD

Hearing dance: poetic or neutral AD?

María Luján Rubio, Marina Ramos Caro, Ana María Rojo López
maria.lujanr@um.es
University of Murcia, Spain

Recent developments in Cognitive Translation and Interpreting Studies have boosted experimental research in Audio Description (AD), highlighting the cognitive processes involved in its AD creation (e.g., Jankowska, 2021; Ramos and Rojo, 2020) and its reception by visually impaired users (Eardley-Weaver, 2013). However, while experimental research has prevailed in AD for film, studies on AD in the performing arts have been mainly descriptive (Fryer, 2018). This study addresses this gap through an experiment assessing the impact of different styles of DP (descriptive vs. creative) on the cognitive and emotional reception of contemporary dance.

Thirty-three visually impaired participants listened to four 5-minute contemporary dance pieces, with descriptive and creative versions of the AD, using an intrasubject design to control for interindividual variability. Reception of the AD was measured using self-report and physiological methods, including cardiac deceleration, skin conductance (EDA), and questionnaires on cognitive effort, liking, emotional arousal, presence, engagement and satisfaction, as well as the ability to recall details of the AD. The AD was recorded using a synthetic female voice and, after listening to the last clip, an open-ended question was included to find out participants' impressions on the voice used.

Results revealed a stated preference for the creative version in retrospective interviews. However, physiological and self-report measures indicated a negative experience with this version. This discrepancy suggests that familiarity with different styles of AD plays a significant role in user perception and satisfaction, highlighting the need for education and exposure to various styles of AD to improve user experience. Furthermore, 95% of participants detected that it was not a human voice, and 20% of participants confirmed that the use of synthetic voice had prevented them from enjoying the experience more, as the voices did not convey and were not in tune with the dance pieces.

This study aims to contribute to the understanding of the interaction between AD styles and their reception by visually impaired audiences in the context of the performing arts, offering valuable insights for

the optimisation of AD practices that promote accessibility and enjoyment for these audiences.

María Luján Rubio is a PhD student in Translation and Interpreting and a substitute lecturer in Translation and Interpreting at the University of Murcia. She also has advanced studies in lyrical singing and is a teacher in secondary education.

Marina Ramos Caro is a Professor of Translation and Interpreting at the University of Murcia (Spain). As a researcher she is particularly interested in the influence of emotions and personality on the process and reception of audio-description and translation. She has received several international grants and has presented her work in a number of prestigious journals, as well as at some thirty conferences around the world. As a professional translator, she has worked mainly in the fields of subtitling and transcreation from English, German and Portuguese into Spanish. She is the founder of Latrium, the Inclusive Translation Laboratory of the University of Murcia, a project dedicated to accessibility in the performing arts for people with sensory disabilities.

Ana Mª Rojo López is a Professor of Translation and Interpreting at UMU. She coordinates the Translation, Didactics and Cognition research group and is a member of the Thematic Network in Empirical and Experimental Research and of the international Laboratory of Cognition and Multilectal Mediated Communication. His research is in the areas of translation and cognition, with a focus on research methodology and the role of emotions in the translation and audio-description process. She has published numerous books and academic papers in prestigious national and international publishers and journals.

Watching and Listening to Dance

Arianna Carloni, Sabine Braun, Dimitris Asimakoulas
a.carloni@surrey.ac.uk
University of Surrey

While audio description (AD) is increasingly used across the creative industries, most research has focused on AD for television and film (Perego, 2018). As noted by Barnés-Castaño et al. (2021), AD for dance remains notably underexplored, despite its growing presence in live as well as online performances. With the growing emphasis on diversity and inclusion in all fields, research aimed at expanding our understanding of dance AD is timely and can contribute to increasing both the demand for and availability of dance AD.

This presentation reports on the findings of a PhD project, first presented at ARSAD 2023, that addresses the research gap in dance AD through a two-part study.

The first part of the project focused on the linguistic aspects of current dance AD, seeking to identify recurring strategies, patterns, and characteristics. A specialised multimodal corpus of dance AD was compiled for this purpose, drawn from described feature films, online clips, and AD scripts from past performances - collected through desktop research and existing archives. The corpus was analysed lexically and semantically, with key features identified, tagged, and annotated to systematise current descriptive strategies.

The second part of the project explored how audiences perceive dance in its visual and in its audio described form. Special attention was given to the unique ways in which individuals construct meaning, engage with, and enjoy the performance - all aspects that have been explored in dance studies (Reason and Reynolds, 2010; Butterworth, 2012 among others). Sighted and blind/partially sighted participants were shown identical dance clips without and with AD respectively, asked to voice their spontaneous reactions through a think-aloud process, and invited to answer some set questions. A thematic analysis of the collected data revealed interesting insights into participants' experience of dance and of its AD. Some findings, such as the critical importance of clarity of expression, were expected; whilst surprising insights emerged regarding how participants derived meaning from dance.

The presentation will give an overview of the rationale and methodological approaches adopted in this project. It will then hone in on key findings around two of the identified linguistic features, namely the use of figurative language and of dance-specific terminology, and their reception among blind/partially sighted

participants. It will conclude by offering some reflections on how the participants' qualitative feedback can be used to guide the selection and effective application of linguistic features in future dance AD.

References

Barnés-Castaño, C., Bernstorff, L., & Vilches, C. (2021). Action research in motion for dance audio description. *New Voices in Translation Studies*, 25.

Butterworth, J. (2012). *Dance studies: The basics*. Routledge.

Perego, E. (2018). Audio description: Evolving recommendations for usable, effective, and enjoyable practices. In L. Pérez-González (Ed.), *The Routledge handbook of audiovisual translation* (pp. 114–129). Routledge.

Reason, M., & Reynolds, D. (2010). Kinesthesia, empathy, and related pleasures: An inquiry into audiences' experiences of watching dance. *Dance Research Journal*, 42(2), 49–75.

Arianna Carloni is a postgraduate research student at the University of Surrey, UK (School of Literature and Languages). In 2013 she completed a Masters in Audiovisual Translation at City University (London, UK) with a dissertation on audio description for dance. After working in the audiovisual translation industry both as a project manager and as a subtitler, she changed careers to pursue her passion for dance and started working for inclusive dance education charities. This rekindled her interest in dance accessibility and led her to start a PhD to explore the topic of audio description for dance in further depth.

PANEL 9. DESCRIBING DIVERSITY

What can I say? Describing physical aspects of people in times of inclusion and diversity

Bernd Benecke
bernd.benecke@br.de
BR Munich, Germany

The description of aspects of a protagonist's physical appearance is a standard in describing fiction or documentary short and feature films. And while things like „blond hair", „grey beard" or „blue eyes" are quite easy to handle and therefore part of nearly every description, things get a little tougher when it comes to the body as a whole.

Is someone short or tall? Is it nice to say, that somebody is slim or skinny or chubby or overweight? Is this just description or interpretation or finally discrimination?

And where do we get when we reach the field of skin colours, ethnic origins or people who define themselves as non-binary?

Looking at this from a historical viewpoint we may say that in the early years of description nearly nobody really cared. Descriptions included everything mentioned above - and it seems that there was no empathy if a description discriminated a person's look, origin, or e.g. sexual orientation. And to be fair: Very often the films itself didn't have this empathy either: Bad characters in movies mainly had dark skin or Southern European or Asian features, gay people were funny with colourful clothing and a weird behaviour. Yes, films in older times discriminated and so did the description.

Then the attention for discrimination grew, films changed (not all) and description had to face the challenge of being non-discriminatory.

The first aspect where this had consequences was the handling with skin colours. For years in Europe „white" was a kind of default and when someone had a different colour this became a part of the description, making it somehow special and very often negative. Up to date blind and visually impaired users demand the skin colour of a person, meaning they want to know when someone is not the classic West European white guy.

Other aspects are the ethnic origins and how to deal with a person that is non-binary? Can we use pronouns and when yes which one? How do we describe physical features without a male or female background?

All this seems so complicated for some companies that they decided to have guidelines that don't allow the description of skin colour, physical aspects like skinny or chubby or anything else that could end in a discussion about discrimination. This in some cases will spoil your description because important things may not be mentioned, and blind users will complain about censorship because something is hidden from them that the sighted audience gets.

This short presentation deals with problems in describing aspects of someone's physical appearance and will give short examples from films where they appear. This is to discuss with the audience, talk about different solutions and decide if it fails or if we all find a satisfying solution.

Bernd Benecke is the Head of Audio-Description at Bayerischer Rundfunk BR – one of the main producers of Audio-Description in Germany. Bernd Benecke started with Description in 1989 when the first movie was described in Germany. He worked as an Audio Description author and narrator for movie distributors and different TV-stations. Today he mainly trains the writers, revises Audio Descriptions and directs the narrating and sound mixing process. He is one of the authors of the German Audio Description guidelines and headed workshops in many countries. Bernd holds a PhD in Translation Studies, published as "Audio Description as partial translation".

Avoiding Sexism, Racism, and Ableism in the Cantonese Audio Description of *Still Human* (淪落人)

Amanda Hiu Tung Chow, Jackie Xiu Yan
amandchow2-c@my.cityu.edu.hk | ctjackie@cityu.edu.hk
City University of Hong Kong

Audio description (AD) renders visual materials to verbal words. It provides the language service that allows people who are visually impaired to grasp more information and interact with diverse groups. AD is also used as an accessibility tool to promote equity and inclusion in society. However, ableist, racist, and sexist language insinuated in human-generated AD and even polarised in artificial intelligence generated AD. In order to improve representations of diversity in AD, researchers have studied how AD intersectionally depicts gender, sexuality, race, and disability. These studies on AD can be found mostly in English and Spanish, rarely in minority languages such as Cantonese. Different languages have their respective strengths and limitations, which may inspire solutions to more equitable and more inclusive expressions in AD and enhance diversity awareness. Therefore, this study examines the representations of gender, sexuality, race, and disability in the Cantonese AD of a Hong Kong movie, *Still Human* (*淪落人*), recounting the relationship between Cheong-wing, a Hong Kong man with paralysis, and Evelyn, Cheong-wing's Filipino caregiver.

The study is conducted first by transcribing the movie AD and then importing the transcription to NVivo for thematic analysis. The words or phrases related to gender, sexuality, race, and disability are tagged and analysed to compile the biases in Cantonese AD, to check if they elicit unnecessary negative sentiments, and to find ways to improve the inappropriate AD. The findings show that the AD carries sexist language in which women are depicted as subordinating to their husbands' household. Although the AD does not contain explicit racist language, only white characters are presented through their skin colour, while the race of non-white characters is primarily indicated by their names or explicitly mentioning their ethnicity. The AD about people with disabilities carries ableist connotations, such as describing their actions through unnecessary negation marking out their functional limitations. Another example of ableist language in the AD can be those highlighting how people with disabilities depend on their caregivers. The Cantonese AD tends to overuse the Chinese character '幫 (to help)'. It is suggested that instead of spelling out how people with disabilities are 'helped' by their caregivers, the description may focus on the actions of the person with disabilities and those of their caregivers to highlight their interdependence.

(This project is partially supported by the GRF project CityU #11609621 and the project #9229103.)

Amanda Hiu Tung Chow is a doctoral student at the Department of Linguistics and Translation, City University of Hong Kong (CityU). She received her master's degree in Comparative Literature from the University of Edinburgh and her bachelor's degree from CityU. Her research interests include audio description and disability studies.

Prof Jackie Xiu Yan received her PhD from the University of Texas at Austin, USA. She is now Subject Leader of the MA Translation Program in the Department of Linguistics and Translation at City University of Hong Kong. Her research and teaching interests include interpreter and translator training, audio description, Chinese culture and translation, and Applied Linguistics. She is now President of the Hong Kong Audio Description Studies Association (HKADSA), President of the Hong Kong Association of University Women (HKAUW). In 2021, she has won a silver medal in the International Exhibition of Inventions of Geneva.

PANEL 10. TRAINING AND CERTIFICATION

AD in Australia: describer accreditation

Karen Seeley
kari.seeley@adelaide.edu.au
The University of Adelaide

This paper proposes developing a professional audio describer identity through a formal accreditation process, driven by legislative reform.

Legislation drives the development of access services (Kubitschke et al, 2013), and it was legislation reform in Australia in the 1970s that resulted in the implementation of world-leading anti-discrimination legislation. The laws mandated the provision of sign language interpreting and/or captioning, to provide the D/deaf and hard-of-hearing community equal access to education, employment, and entertainment (Bancroft, 2018, Pöchhacker, 2015). Legislation became the impetus for developing the professional sign language sector in Australia, driving accreditation for sign language interpreters, delivered through the National Accreditation Authority for Translators and Interpreters (NAATI) (Bell, 1997). In turn, that accreditation was foundational for the development of the professional sign language interpreter identity, with the consequent development of professional standards and pay rates.

However, Australia's anti-discrimination legislation does not mandate access provisions, such as AD, for those in the b/vi community (Seeley 2023). Therefore, legislative reform is key to redressing this inequity and driving the development of the professional audio description sector. Accreditation is one of the elements required in the process of developing the professional describer identity, and mandated legislative compliance would drive the development of a pathway to professional describer accreditation, along with the development of professional standards and pay rates.

This paper proposes developing a professional audio describer identity through a formal accreditation process, driven by legislative reform. However, broadening the understanding of interpreter cohorts to include audio describers will need to be underpinned by a commitment to cooperation and collaboration between the academy, practice, and an accreditation association (such as NAATI in Australia) if the proposed approach to describer accreditation is to be adopted and result in improved outcomes for end-users (Seeley 2025). The proposed approach to developing an Australian audio describer

identity through accreditation driven by legislative reform may be applicable in other jurisdictions.

References

Bancroft, M. A. (2018). Community interpreting: A profession rooted in social justice. In H. Mikkelson & R. Jourdenais (Eds.), *The Routledge handbook of interpreting*. London and New York: Routledge.

Bell, S. J. (1997). The challenges of setting and monitoring the standards of community interpreting: An Australian perspective. In *The critical link: Interpreters in the community*. John Benjamins.

Kubitschke, L., Cullen, K., Dolphi, C., Laurin, S., & Cederbom, A. (2013). *Study on assessing and promoting e-accessibility.*
Pöchhacker, F. (2015). Evolution of interpreting research. In *The Routledge handbook of interpreting.* London and New York: Routledge.

Seeley, K. (2023). *End-user experiences of audio description for live theatre: Complexities beyond practice* (PhD thesis). The University of Adelaide.

Seeley, K. (2025). Audio description and interpreting studies: Complexities beyond practice. In C. Zhan & R. Moratto (Eds.), *Audio description and interpreting studies: Interdisciplinary crossroads*. London and New York: Routledge.

Karen Seeley, Adjunct Fellow in the School of Humanities at The University of Adelaide. She holds a BA in Politics and International Studies, a BA (Hons) First Class in Media, and a PhD in Audio Description. She trained as a professional Audio Describer in both Australia and the US and has been working as a freelance audio describer since 2014, in a broad range of settings including live performance and digital format. She has delivered professional AD training in Australia and Singapore. She has published chapters about AD in Australia and the intersection of AD and Interpreting Studies. She uses an interdisciplinary framework through which to focus on the end-user experience of AD. Acknowledging end-users as experts in their own lives, her research seeks to understand the impact of broader contexts, beyond AD practice, which shape the experience of using AD services to access social and cultural participation.

Audio description: a transversal training tool for minoritized language education

Floriane Bardini, Eva Espasa
floriane.bardini@uvic.cat
Universitat de Vic

In the past few years, research has increased on didactic audiovisual translation, including traditional audiovisual translation modalities such as dubbing or subtitling, as well as accessibility practices such as audio description (AD) and subtitling for the deaf and hard-of-hearing. Such research has mostly focussed on foreign language teaching, but audiovisual accessibility and translation (AVAT), and especially audio description, can also be successfully applied to minoritized language teaching.

These applications are the object of the Erasmus+ HOW TIMELY project (KA220-SCH-6F8B23D3, 2025-2027), which promotes Minority and Endangered Languages (MEL) revitalization among young users through the practice of AVAT. It focuses on three MEL (Catalan, Irish, Sardinian) and 3 AVAT modes (subtitling, re-voicing, and AD). The main output of the project will be the open access HOW TIMELY teaching/training portal, which will provide self-education resources for MEL teachers and AVAT activities to implement in class. In our presentation, we will focus on the approach to didactic audio description (DAD) adopted in the project.

To promote the use of minoritized language in the minoritized settings of the project partner schools, DAD activities will be created. Strategies will also be developed to foster the instrumental use of the minoritized language for an AD-oriented film analysis, and for reflection on the film's content and on the AD task itself. Before developing these activities and strategies, a survey will be conducted among MEL teachers to enquire about their prior knowledge of audiovisual accessibility and about their needs to implement DAD activities in class.

We will present a set of guidelines based on current professional AD guidelines and adapted to DAD, which can be fine-tuned according to the level of tuition. These guidelines will help students draft authentic AD scripts, while they also will help teachers to evaluate the tasks. We will give insights into how DAD activities in secondary school can enhance students' linguistic skills according to the CEFR (2020): reception, production, interaction, and mediation. DAD fosters lexical and syntactic skills, as well as aural abilities, but it also helps to develop transversal competences such as digital, interpersonal, and strategic competences.

Finally, focussing on AD triggers awareness on inclusiveness for sensory disability. Therefore, an important expected outcome of the project is the engagement of teachers, learners, and stakeholders in awareness on linguistic and sensory accessibility, along with MEL revitalization across the EU, especially among the youth.

References

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume.* Council of Europe Publishing. Retrieved from https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Dr Floriane Bardini is an associate lecturer at the University Pompeu Fabra and at the University of Vic - Central University of Catalonia (UVic-UCC), where she conducts research in the field of audiovisual translation (AVT) and media accessibility as a member of the research group TRACTE (Audiovisual Translation, Communication and Place, SGR 2021 SGR 0814). Her main research interests include audio description, user experience, AVT and minority language teaching. Bardini is a member of the board of Linguapax International, an NGO dedicated to the protection and revitalization of world linguistic diversity, and of AccessCat, the Accessibility Innovation Network of Catalonia which promotes access to information and communication.
ORCID: 0000-0001-5116-6631

Dr Eva Espasa is full professor in Translation Studies at the University of Vic - Central University of Catalonia (UVic-UCC). She lectures in audiovisual translation and translation for advertising at undergraduate and postgraduate levels, mostly from English into Catalan. Her main research interests are audiovisual translation and accessibility, theatre translation and gender studies, topics on which she has lectured and published extensively. Espasa is coordinator of the interdisciplinary research group TRACTE (Audiovisual Translation, Communication and Place, SGR 2021 SGR 0814). She is a member of AccessCat, an Accessibility Innovation Network of Catalonia promoting access to information and communication.
ORCID: 0000-0002-0463-9317

Film students and audio description: results of a francophone inclusive media training pilot project in Québec, Canada

Caroline Martin, Santiago Hidalgo
caroline.martin.6@umontreal.ca
Montreal University

The «Formation en Médias Inclusifs: Audiodescription/Vidéodescription et Sous-Titrage par Reconnaissance Vocale» is the first Francophone initiative of its kind in Canada. This pilot project, conducted from 2022 to 2023 at the Université de Montréal (Quebec, Canada), was funded by the Broadcasting Accessibility Fund. This talk outlines the genesis of the training program, its objectives, curriculum, and the outcomes observed among both teachers and students.

Genesis: The idea of developing a French-language training course in inclusive media emerged from our participation in the advisory committee for the Inclusive Media Course Series at Toronto Metropolitan University (Canada), led by Professor Deborah Fels. This invitation highlighted the potential of film studies to contribute to research on audio description and to the training of audio describers within the film student community.

Objectives: The training had two primary objectives: first, to train audio describers and live subtitlers (using respeaking techniques) to meet the demands of Canadian broadcasters for French-language content. In 2016, the CRTC mandated that, by September 1, 2019, all major broadcasters must provide videodescription for all eligible programs aired between 7 p.m. and 11 p.m. (prime time), seven days a week.

Course curriculum: The pilot training project consisted of three courses: Introduction to Audio and Video Description, Introduction to Voice Recognition Subtitling, and Inclusive Creation: Theory, History, and Practice. The first course provided an overview of the history of audio description and studies on its reception, with a focus on developing technical skills such as scripting, narration, recording, and quality control.

Results: Feedback from both teachers and students affirmed the importance of such training within the undergraduate film studies curriculum. In addition to acquiring technical skills, which students viewed as an asset for entering the job market, they emphasized the need to incorporate such training into the core curriculum.

References

CRTC. (2016, November 2). *Broadcasting Regulatory Policy CRTC 2016-436*. Canadian Radio-television and Telecommunications Commission. https://crtc.gc.ca/eng/archive/2016/2016-436.htm

The Chang School of Continuing Education, Toronto Metropolitan University. (n.d.). *Course series inclusive media: Real-time closed captioning and audio description/described video*. https://continuing.torontomu.ca/public/category/courseCategoryCertificateProfile.do?method=load&certificateId=281328

Université de Montréal. (n.d.). *Formation en médias inclusifs: audiodescription/vidéodescription et sous-titrage par reconnaissance vocale*. https://audiodescription.labocinemedias.ca/

Caroline Martin has taught literature and film for over 20 years Unviersité de Montréal and at the Université du Québec en Abitibi-Témiscamingue.
Her work focuses on film education, reception among young audiences and audio description. She is co-editor of the special issue "Éducation cinématographique: trajectoires France-Québec" in Nouvelles Vues. Revue sur les pratiques, les théories et l'histoire du cinéma au Québec. She is also the principal researcher and author of an article on the CINÉPROF research project, "L'enseignement du cinéma dans les écoles secondaires québécoises: portrait d'un curriculum enseigné inspiré des arts plastiques et de la littérature", in the Revue des sciences de l'éducation. Recently, she participated in the Inclusive Media Training program at the Montreal University, both as a content co-developer and as a teacher.

Santiago Hidalgo is an assistant professor in the Department of Art History and Film Studies at the Université de Montréal. He is the director of Laboratoire CinéMédias (since 2018) and executive director of the cinEXmedia partnership, « Ensuring Well-Being in the Screen Age », focusing on the societal, pedagogical and therapeutic challenges of the cinematic experience from an inclusive perspective (2022-2029). In 2019, Hidalgo co-founded the Cinema and technology collection at Amsterdam University Press, as well as Cinéma et technologie at University Press in 2021. He co-directs both collections. He is currently a researcher at the Centre de recherche de l'Institut universitaire de gériatrie de Montréal and the Centre d'études avancées en médecine du sommeil, and co-leader of the Inclusive Media Training program at the Montreal University.

'Accessibility is Like Air - Everyone Needs it to Stay Alive' - The Impact of Accessibility Training on Awareness Levels

Anna Jankowska, Nina Reviers, Gert Vercauteren
anna.jankowska@uantwerpen.be
University of Antwerp

Accessibility is critical for enabling the independent living and full participation of individuals with disabilities. As Greco (2019) highlights, recent years have witnessed an "accessibility revolution" in research, legislation, and service provision. However, while legal mandates are essential, the full potential of this revolution can only be realized when those responsible for implementing these requirements are driven by awareness and sensitivity rather than mere compliance.

Our presentation discusses the findings of a survey-based study conducted in Poland and Belgium, which aimed to assess university students' awareness of accessibility. The study specifically compared students who had received media accessibility training with those who had not. The results revealed that students without media accessibility training largely lacked knowledge about the needs of individuals with disabilities and the necessary measures for ensuring accessibility in media. In contrast, even basic training led to a significant increase in awareness and sensitivity towards accessibility issues, along with a broader understanding of how media accessibility benefits not only those with disabilities but also the wider population. One participant poignantly described accessibility as being "as essential as air," emphasizing its fundamental importance for everyone.

Our findings support the argument that incorporating accessibility training into mainstream educational curricula, as recommended by the Young Archers and ATHENA projects, yields positive outcomes. In the long term, this approach can challenge misconceptions and biases about disabilities, foster a more inclusive society, and encourage proactive design and policy decisions. By shaping informed professionals, experts, and citizens, accessibility education can contribute to more effective and sustainable accessibility practices that go beyond legal compliance.

Anna Jankowska (University of Antwerp) is an AVT & MA researcher, trainer and consultant, former audio describer and access manager. She is now a research professor at of University of Antwerp. She is the founder and was the CEO of the Seventh Sense Foundation which provided access services in Poland. Her recent research projects include studies on audio description process, mobile accessibility and software.

Nina Reviers (University of Antwerp) is a tenure-track professors in Audiovisual Translation and Media Accessibility at the Department of Applied Linguistics, Translators and Interpreters. She holds a PhD in Translation Studies (University of Antwerp, 2018) in the field of Audio Description, for which she was awarded the EST Young Scholar Prize in 2019. Her research addresses linguistic and multimodal aspects of audio description, computer-aided translation of audio description, integrated access for the (scenic) arts and technology for access. As manager of the OPEN Expertise Centre for Accessible Media and Culture, Nina fosters a close collaboration with stakeholders as a key factor in her research and teaching activities.

Gert Vercauteren is a tenure track professor at the University of Antwerp where he teaches translation technology and audiovisual translation/communication. He holds a PhD in Translation Studies (University of Antwerp), more specifically in the field of Audio Description. His current research focuses on narratology in audio description with a particular focus on sound and on cognitive load in/and audio description. He also works as an audio describer, maintaining a close link with the audiovisual translation industry.

PANEL 11. COGNITIVE ACCESSIBILITY AND AD

Accessible Screenings for Older People: Enhancing Cognitive and Social Well-being through the FAIR Project

Francesca Carpenedo, Federico Spoletti
fair@subti.com
Sub-Ti

The FAIR project aims to promote the enjoyment of enhanced film content among older people, particularly those with sensory disabilities and those residing in remote areas, who are often excluded from cultural participation. This presentation will detail the innovative approach of the FAIR project, which utilizes advanced technology to make films accessible via subtitles for the deaf and hard of hearing and audio description for the blind and visually impaired. The project primarily targets older people residents in nursing homes but also benefits all sensory disabled persons, enhancing their cultural inclusion and digital literacy.

Key features of the project include the development of user-friendly controls and the ability to view films on various devices, ensuring ease of use even for those with limited digital skills. The accessible content can be enjoyed both individually and collectively, with each person able to customize their viewing experience via personal devices without disturbing others. This approach is especially suitable for nursing homes and similar settings, where collective viewing can foster social interaction and community engagement.

The FAIR project also addresses broader issues of user involvement and ethical aspects by promoting technological transition and inclusion for people with disabilities. Training by facilitators ensures that users can effectively utilize the technology, advancing their digital competence.

Additionally, the presentation will highlight the cognitive and social benefits of accessible screenings, which may help slow down dementia and cognitive decline by keeping the brain active. Participation in cultural activities has been shown to combat loneliness, reduce stress, and improve mental and physical well-being, contributing to a higher quality of life for older people.

This session will also align with conference topics such as "Audio Description Experiences Across the Globe in Diverse Formats and Environments" and "User Involvement and Ethical Aspects," showcasing how the FAIR project can be a model for similar initiatives internationally. By expanding this project abroad, we aim to create an

inclusive cultural heritage accessible to anyone with sensory disabilities, ensuring no audience is left behind.

Francesca Carpenedo, a graduate in Economics from the University of Trieste, has been dedicated to social issues concerning older people since the start of her career. In 1998, she co-founded a cooperative providing non-residential services for non-self-sufficient older individuals, overseeing its organization and management. In 2001, she launched the local branch of Telefono Anziani Maltrattati (TAM), aimed at collecting and managing reports of abuse against older people, which quickly became a national reference point.
In 2014, through TAM, Carpenedo joined AGE Platform Europe, a European network promoting the interests of older adults and raising awareness of issues affecting them. In 2024, she was appointed to its Board of Directors. She also participated in the WHO's Healthy Ageing course in 2020 and, in 2023-24, contributed to the FAIR project, which promotes accessible cultural content in residential facilities.

Born and raised in Italy, **Federico Spoletti** is the co-founder and Managing Director of SUB-TI, an international subtitling company based in London, providing audiovisual translation services worldwide. He also co-founded SUB-TI ACCESS, a company dedicated to cultural accessibility for the hearing and visually impaired, specialising in subtitles for the deaf and audio description for the blind.
In 2011, Federico launched FRED Film Radio - The Festival Insider, an internet radio network with 29 channels broadcasting in 25 languages, including thematic channels focused on film education and the film industry.
In 2023, he launched INCinema, the first film festival fully accessible to people with sensory disabilities. He has been involved in numerous international projects, including FRED at School, LTA, SMART, WICIP, and more recently, FAIR (accessible screenings for the elderly in care homes).

Repurposing Audio Description, Audio Narrative and integrated subtitles: trialling 'Accessible Cues' with cognitively diverse audiences in the UK

Maria Andreea Deleanu, Constantin Orasan, Sabine Braun
m.deleanu@surrey.ac.uk
University of Surrey

This presentation explores the potential of enhancing the cognitive accessibility of audiovisual narratives for diverse audiences, particularly those with cognitive impairments, by expanding on existing Audio Description practice. Since the advent of moving pictures, the way we tell stories and consume storyworlds (Herman, 2002) has evolved into a multimodal language which can be difficult for some audiences to navigate. Despite significant advances in Audiovisual Translation and Media Accessibility in recent years (European Parliament, 2007), the needs of cognitively diverse individuals have been largely overlooked. While institutions (European Commission, 2019, 2021; European Parliament, 2016) have started offering domain-specific content into Easy-to-Understand (E2U) language to prevent communicative exclusion of those facing cognitive barriers (UNCRPD, 2006; COGA, 2017; Maaß, 2020; Perego, 2020), the effort has primarily focused on Web and public-sector communication (European Commission, 2016), with no established practice for multimodal formats like films.

Films are not just entertainment; they promote inclusion, participation and education by providing insight into emotions and behaviors (Thonon et al., 2016; Wassmann, 2015). Yet, processing complex audiovisual narratives such as films and TV productions can prove challenging for people with diverse cognitive abilities (APA; 2013; Schlickers, 2009), potentially excluding them from fully comprehending and enjoying these narratives.

The project presented in this paper addresses the cognitive divide in Media Accessibility by developing a conceptual solution for the integration of explanatory cues in (cognitively inaccessible) films, called Accessible Cues, and trial it with cognitively diverse audiences. These cues combine features of Audio Description and Audio Narration (Kruger, 2010) with other elements that can aid audience comprehension. The development of the conceptual solution was based on research into film comprehension challenges for cognitively diverse audiences and research into E2U, including a comprehensive analysis of E2U guidelines (Deleanu et al., 2024). To evaluate the effectiveness of this solution, prototype Accessible Cues were embedded in an abridged version of a romantic comedy, enriching it with narration, captions and visual prompts provided in E2U format. The prototype was then trialled with n=76 participants recruited

across several user associations in the UK. The study involved a series of questionnaires to elicit demographic information, viewing habits and experiences to better understand the sample. The core of the study was, however, the film viewing experience, as well as the participants' level of comprehension and enjoyment of the film and their perceptions of Accessible Cues. The methods and findings from the reception study will be illustrated during the presentation, highlighting the experiences and performance of participants. The findings suggest that Accessible Cues improve comprehension and enjoyment of audiovisual narratives for the cognitively diverse users in our sample. Based on this, a broader adoption of Accessible Cues in films and other types of audiovisual content could provide several cultural and educational benefits to cognitively diverse audiences.

The presentation will focus on the reception study conducted within the project. It will begin by outlining and critically reflecting on the methodological approach and instruments used for the reception and then present the key findings, highlighting the experiences and performance of participants. To conclude, the presentation will discuss the implications of findings for practice and further research.

Andreea Deleanu is a PhD Candidate at the University of Surrey, UK, specializing in Media Accessibility under the supervision of Sabine Braun and Constantin Orasan. Her project focuses on repurposing Easy to Understand language, Audio Description, Audio Narration and integrated subtitles for audiences with diverse cognitive abilities, to guide and support their comprehension, enjoyment and immersion when accessing audiovisual narrative.
She holds an MA in Conference Interpreting from the University of Trieste, Italy and has specialized as a healthcare interpreter and audiovisual translator.
Andreea was a Visiting Scholar at Macquarie University, Australia, where she developed her 'Accessible Cues' under the guidance of Prof Jan-Louis Kruger.
Andreea was also involved in the EU project EASIT (Easy Access for Social Inclusion Training) under the guidance of Prof Elisa Perego. She focused on unit 3, on the application of E2U language to AD to improve access for users with visual impairment. More information about the projects and its outputs can be found here: https://transmediacatalonia.uab.cat/easit/

Dr Sabine Braun is Professor of Translation Studies and Director of the Centre for Translation Studies at the University of Surrey, UK, a Research-England funded Centre of Excellence. She also serves as a Co-Director of Surrey's Institute for People-Centred AI. She specialises in human-machine integration in translation and interpreting to improve access to information, digital content and public services. For over a decade, she spearheaded a European-

funded research programme investigating video-mediated interpreting in legal proceedings to improve language access in the justice sector (AVIDICUS 1-3; 2008-16), while contributing her expertise in video interpreting to other justice sector projects (e.g. QUALITAS, 2012-14; Understanding Justice, 2013-16; VEJ Evaluation, 2018-20). Subsequently she advised justice sector institutions on the use and risks of video-mediated interpreting, delivered training, developed European guidelines, and co-authored a DIN standard. She has also explored the use of video and virtual reality platforms for training interpreters and users of interpreting services (IVY, 2011-3; EVIVA, 2014-15; SHIFT, 2015-18; EU-WEBPSI, 2021-24) and is currently involved in projects investigating the application of communication technologies and AI-enabled language technologies to multilingual health communication (MHealth4All, 2021-24; Interpret-X, 2021-24). Furthermore, she conducts research on audio description and other translation modalities related to accessible communication. In the Horizon 2020 project MeMAD (2018-21), she explored the potential for (semi-)automating audio description to enhance digital media inclusion. In 2024, she launched a Leverhulme Trust-funded Doctoral Training Network on AI-Enabled Digital Accessibility (ADA). Her overarching interest centres on fairness, transparency, and quality in the use of technology in language mediation.

Constantin Orasan is Professor of Language and Translation Technologies at the Centre of Translation Studies, University of Surrey, UK and a Fellow of the Surrey Institute for People-Centred Artificial Intelligence. Before starting this role, he was Reader in Computational Linguistics at the University of Wolverhampton, UK, and the deputy head of the Research Group in Computational Linguistics at the same university. He has over 25 years of experience in the fields of Natural Language Processing (NLP), Translation Technologies, Artificial Intelligence and Machine Learning for language processing. His recent research focuses on the use of Generative AI as a support tool for translators and the use of Automatic Speech Recognition for interpreters. In the past he was the deputy coordinator of the FIRST project, a project which developed language technologies for making texts more accessible to people with autism. In addition to managing a consortium of nine partners from academia, industry and heath care organisations, he also carried out research on text simplification and contributed to the development of a powerful editor which can be used by carers of people with autism to make texts more accessible for these people. His research is well known in these fields as a result of over 130 peer-reviewed articles in journals, books and international conferences. More information about him can be found at https://dinel.org.uk/