Audio description and audio subtitling

Elisa Perego (University of Pavia)

ENACT project workshop – 19 march 2025

We know that the action of verbally assisting blind people with descriptions of the things around them is not new: audio description has always existed in some form, albeit rudimentary. But it was only in the 1970s that audio description began to be systematically addressed — both as a service and as a discipline.

Since then, **we have reached significant milestones**.

One of them is the fact that **we can no longer call AD a "new" discipline** — it has firmly established itself || as an essential part of media accessibility. Another milestone is the **10th anniversary of ARSAD**, which has now been around for 20 years.

To examine the evolution of AD, I decided to compare the ARSAD programmes from 2009 (the oldest I could find) and from 2025. A comparison that gives a clear picture of how the field has progressed.

In 2009, the focus was on raising awareness, defining best practices and the belief that AD is an essential service and not an optional add-on. Discussions revolved around the training of describers, and how AD can be extended beyond film and television.

The landscape has changed dramatically. We now discuss personalization, artificial intelligence, automation, machine-generated scripts and AI-powered accessibility. The field has shifted from simply providing access to rethinking the entire user experience.

One of the most profound changes has been in **technology**. In 2009, we relied primarily on human describers and voice talents, we struggled with translation challenges, and worked to expand market growth. Today, AI is everywhere. We have machine learning models capable of generating descriptions in real time, AI-powered translation tools and generative AI creating AD for films, video games and museums, and also highly improved synthetic voices. **The role of the human describer is evolving — from**

creator to curator, ensuring that AI-generated descriptions retain quality, accuracy and emotional depth. And this has to do with **language** as well: people learn language also from the media they consume. This means that AD carries with it a huge responsibility. This means that we cannot allow inferior language in AD, as this risks degrading language skills rather than enriching them, even in contexts where simplified forms of AD are or have to be used.

Another major shift is the **expansion of media:** Today, accessibility extends to video games, virtual reality, interactive storytelling and AI-driven museum guides. We are no longer just making existing content accessible, but integrating accessibility into the structure of media creation ensuring *immersive* and *inclusive* experiences.

We have also seen a fundamental shift in the way we think about the **user**. In 2009, the focus was on 'correct' descriptions - correct terminology, clarity and adherence to guidelines and timelines. Today, it's no longer *just* about what we say, but how users perceive and experience AD. Research is looking at cognitive accessibility, investigating how AD affects cognition, and even using

biometric tools to understand engagement and emotional responses.

And despite these advances, one crucial aspect remains **underdeveloped: the inclusion of end users in the AD chain**: true co-creation and user-centred design are still not the norm, but should become a priority.

We are also dealing with profound **ethical issues**. How do we ensure that AI-generated descriptions remain neutral and unbiased? How do we represent diversity in AD? How do we strike a balance between automation and the irreplaceable human touch? These were once theoretical questions, but today they are at the centre of our discussions.

The landscape of **education and training** has also evolved. In 2009, university courses on AD were just emerging. Today, there are global standards, certification programmes and interdisciplinary collaboration. AD is no longer a niche skill, but an integral part of media and language education.

In addition, the field of accessibility has expanded to include **minority languages**: AD is increasingly offered in diverse language communities and also on several streaming platforms. This demonstrates the **growing recognition of AD** as such, **and of AD professionals** (be them describers or QC-ers), particularly in countries where they were not previously recognised.

This tells us that we have come a long way. The first ARSAD seminars laid the foundation, but today's ARSAD shows what comes next: a **future** where **technology and human expertise** work together.

As we look to the AD future changes, we must lead them, push the boundaries of what is possible, and never lose sight of the human element. Technology is a tool, but I believe the inner quality of audio description will continue to be determined by human expertise, empathy and creativity, and by the enthusiasm and the talent of all the people involved.