


---

This is the **accepted version** of the book part:

Matamala, Anna. «Audio description». *Handbook of Translation Studies*, Vol. 5 (2021), p. 17-22 DOI 10.1075/hts.5.aud2

---

This version is available at <https://ddd.uab.cat/record/327626>

under the terms of the  BY COPYRIGHT license.

## **Audio description**

*Anna Matamala*

An audio description (AD) converts images into words: visual elements—and sounds which cannot be understood without the images—are transferred into AD units, i.e. aural linguistic elements which are inserted when there is no significant speech or relevant sound in the source content (the so-called “silent gaps”).

The combination of the original soundtrack with the AD creates a new coherent whole that provides access to those who cannot properly see the visuals. This includes persons with sight loss, but also sighted persons who, for a wide variety of reasons, may not have access to the visuals or may want to enjoy them with AD.

AD can be provided for audiovisual and visual content such as films, TV programmes, opera, dance, musical and theatre performances, sports events, art works, natural and cultural heritage sites, live events, among many others. When subtitles are present in a content, AD is offered together with audio subtitling, an oral rendering of the written subtitles. In some cases, especially in the performing arts, AD can be preceded by an audio introduction and a touch tour. An audio introduction is a recorded or live presentation which generally offers a short synopsis and a description of the main characters and visual style of the play. Touch tours are tactile visits which allow to explore the stage and touch the props.

AD can be said to have originated informally when sighted persons explained friends and family with sight loss what was happening on an audiovisual content. Ancient film

narrators—or *benshi* in Japan—, radio dramas, opera or film radio broadcasts, and even literary figures such as ekphrasis have been considered precursors of AD in the literature. However, the origins of AD as such are generally traced back to the United States in the 1970s, with Frazier’s “television for the blind”, a concept later transferred to a regular service at Arena Stage Theater in Washington by Margaret Pfanstiehl. The literature acknowledges that the first audio described film was broadcast in 1983 by the Japanese channel NTV.

AD can be viewed as an intersemiotic type of audiovisual translation\* or as an access service within the realm of media accessibility\*. The relationship between audiovisual translation and media accessibility is still open to discussion and boundaries are unclear. In fact, as Matamala (2019) puts it, AD can be approached both as an audiovisual transfer model within audiovisual translation and as an access service within what can be termed “audiovisual accessibility”.

A broad overview of AD in terms of practice and research is found in the monographs by Fryer (2016) and Snyder (2014), and also in edited volumes by Maszerowska, Matamala and Orero (2014) and Matamala and Orero (2016). All these works have inspired the present entry.

## 1. AD categorisation and main features

The specific characteristics of the content to be audio described have an impact on the actual AD. The content can:

- range from films (horror, thriller, comedy, drama, etc.) to the performing arts (dance, theatre, circus, etc.), cultural and natural heritage, or live events, among others.
- be an original production or a translation produced through dubbing\*, voice-over\*, subtitling\* or other transfer modes.
- be static (for example, a painting) or dynamic (for example, a circus performance).
- be recorded or live.
- be planned, semi-planned or improvised.

The process of AD can also differ in many aspects. The AD can be:

- delivered live or recorded.
- planned, semi-planned or improvised.
- created as part of production or postproduction processes.
- created or translated. In both cases, it can be created or translated by a human or by a machine, with a possible human revision afterwards (post-editing).
- delivered by a human, or by a synthetic voice through a text-to-speech system.
- open or closed, meaning it can be heard by anyone watching the content (open) or can only be activated when the viewer selects it (closed).
- mixed at the broadcaster or server end or mixed at the user end.
- generally delivered in one language, but there are examples in live AD in which two languages understood by the audience are combined also exist.

An interesting debate has been that of objectivity versus subjectivity. Most early guidelines emphasize that describers should describe what they see (WYSWYS, What You See is What You Say), avoiding any type of manipulation or interpretation as this

could be considered patronizing. However, defining what one sees is not so straightforward. We do not all see the same, as our previous experiences and knowledge shape our understanding, and research with users proves that certain audiences are willing to accept more creative approaches (Mazur and Chmiel, 2012). In any case, the time available for each AD unit is generally limited, and describers need to find strategies to select the better approach when deciding (a) what to describe and (b) how to describe it.

Narratology has been considered to provide a sound theoretical framework in relation to what to describe, especially in filmic AD. Remael, Reviers and Vercauteren (2015) identify some central constituents and provide recommendations on how to analyse the source text and how to develop an AD taking into account two main narratological building blocks: characters (and their actions), and spatio-temporal settings.

- Characters—protagonist, antagonist, and supporting characters—and their actions and reactions move the story forward. Characters can be new or known, they can be authentic or fictional, real or unrealistic, and they can be related or unrelated. They can have a symbolic function or help to indicate a lapse of time.
- Spatio-temporal settings refer to the place and time where the action takes place. They can be global or specific, real or imagined, well-known or unfamiliar. They can be presented explicitly or implicitly. They can have a background function or a narrative/symbolic function.

By understanding how the story is constructed, the describer can reconstruct it in the AD. Other aspects that Remael, Reviers and Vercauteren identify as worth paying attention to

are related to film techniques—namely film language, sound effects and music, text on screen—, and intertextual references.

Regarding how to describe it, language is considered a central aspect. Perego (2019) indicates that the adjectives that better describe the language of AD are “meticulous”, “concise”, “visually intense”, and “usable”. In other words, the language of AD is expected to be detailed, accurate and precise, fitting in a limited space and conveying clear images. Plain syntax and a logical organisation of information are advised, to facilitate end user processing.

## 2. The AD process

Four main stages can be identified in the AD process, although some adjustments may need to be made in contexts such as live improvised AD:

- Preparation: the source text is analysed and a documentation process starts, following the instructions received (brief).
- Creation process: a written script is produced and timed.
- Voicing and recording: the script is voiced, either by the same describer or by a voice artist, and recorded. In recorded AD, this step is followed by technical procedures such as the sound mix.
- Quality control: a quality evaluation process, ideally including end users, is undertaken.

There are both local and international standards and recommendations that provide guidance on the process, such as *ISO/IEC TS 20071-21:2015 Information Technology – User interface component accessibility. Part 21: Guidance on audio description or Web Content Accessibility Guidelines (WCAG) 2.1*. Rai, Green and Leen (2010) also present a comparative study of different national standards and recommendations. The ADLAB guidelines (Remael, Reviere and Vercauteren, 2015) merit special attention: rather than advocating for a prescriptive approach, they suggest strategies to deal with different problems during the AD process.

### 3. Research

Research on AD began at the end of the 20th century outside the field of translation, when Gregory Frazier developed an MA dissertation (1975) on the topic. In audiovisual translation, AD has gained momentum with the turn of the century, and the Advanced Research Seminar on Audio Description (ARSAD), held at Universitat Autònoma de Barcelona every two years, has become a conference of reference in the field.

From a theoretical point of view, cognitive (Holsanova), discursive (Braun) and narratological (Kruger, Vercauteren) models have been used to account for the processes involved in AD.

From a descriptive point of view, publications analyse AD in a plethora of genres and languages, sometimes providing an historical overview and a critical analysis of existing legislation. Special emphasis has been put on filmic AD, with investigations dealing with typical research topics in translation studies such as cultural\* references, humour\*,

music\*, voices\*, but also more specific aspects such as how to describe characters, spaces, facial expressions, or written language.

Going beyond specific case studies, different corpora\* have been built to better understand AD: TIWO (Television in Words) at University of Surrey, TRACCE, MOVD (MoMA Visual Descriptions) and MOAP (MoMa Audio Programme) at University of Granada, and VIW (Visuals Into Words) at Universitat Autònoma de Barcelona are examples of such corpora.

Some European projects have included AD as an object of study. The Pear Tree Project (Mazur and Kruger, 2012), developed within the European project DTV4ALL (2008-2010), aimed better to understand film reception across Europe, trying to find cross-linguistic and cross-cultural commonalities and differences as a first step towards developing a common set of AD guidelines. This work evolved in part into the ADLAB project (2011-2014), which mapped practices and guidelines across Europe and finally produced reference guidelines in the field (Remael, Reviers and Vercauteren, 2015).

DTV4ALL, and its successor HBB4ALL (2013-2016), also tested different technical scenarios for the distribution and reception of AD. Whereas the former focused on digital television, the latter dealt with hybrid connected television. A step further has been taken in the ImAc project (2017-2020), in which AD in virtual environments has been developed and evaluated, in terms of technical delivery, sound treatment and scripting.

From a technological point of view, the application of language and translation technologies was researched by the ALST project (2013-2015): more specifically, ALST

investigated whether speech recognition, machine translation (with postediting) and text-to-speech technologies could be integrated into the AD workflow. In addition, computer vision techniques are used to research the automatic generation of audio descriptions (see MeMAD project, 2018-2020). Other projects, such as Open Art in Poland, placed the emphasis on developing and testing apps to provide AD in museums.

The didactics of AD—and AD as an educational tool—have also been a relevant research topic, with special attention paid to the skills and competences audio describers need to acquire. A seminal project was ADLAB PRO (2016-2019), which assessed current AD training practices in Europe, defined the profile of the AD professional, and designed a course and developed educational resources which are now freely available on the internet. Along similar lines, the EASIT project (2018-2021) researches how AD can be merged with Easy-to-read or Plain Language to provide easy-to-understand audio descriptions to diverse users, and aims also to produce open access educational content.

Research in AD has also placed a strong focus on how end users receive different AD strategies (Chmiel and Mazur, 2012). As an example, audio descriptions created by a human narrator have been compared to audio descriptions by a text-to-speech system. Similarly, audio descriptions created with a standard approach have been compared to more creative audio descriptions or to descriptions including cinematic language. And audio descriptions with diverging amounts of information have also been compared. These are just three illustrative examples of extensive reception research, which has gathered users' preferences, comprehension, recall or emotional response, to name a few measures, by using both subjective and objective methods. Persons with sight loss have been central in reception\* research, but studies on AD with sighted viewers have also

taken place, in order to better understand how sighted viewers watch movies and in order to identify additional uses of AD.

Overall, although AD research and practice is relatively recent compared to other modes, the intersemiotic transfer and multimodal interplay that take place in any AD have made it an attractive research topic that can be approached from a myriad of perspectives.

### **Further essential reading**

Fryer, Louise 2016. *An introduction to audio description. A practical guide*. London: Routledge.

Maszerowska, Anna, Matamala, Anna, and Pilar Orero (eds). 2015. *Audio description. New perspectives illustrated*. Amsterdam: John Benjamins.

Matamala, Anna, and Pilar Orero (eds). 2016. *Researching audio description. New approaches*. London: Palgrave Macmillan.

Perego, Elisa. 2019. Audio description: evolving recommendations for usable, effective and enjoyable practices. In *The Routledge Handbook of Audiovisual Translation*, ed. by Luis Pérez-González, 114-129. London: Routledge.

Remael, Aline, Reviers, Nina, and Gert Vercauteren (eds). 2015. *Pictures painted in words. ADLAB audio description guidelines*. Trieste: EUT

## References

Chmiel, Agnieszka, and Iwona Mazur. 2012. AD reception research: Some methodological considerations. In *Emerging topics in translation: audio description*, ed. by Elisa Perego, 57-80. Trieste: EUT.

Fryer, Louise 2016. *An introduction to audio description. A practical guide*. London: Routledge.

Maszerowska, Anna, Matamala, Anna, and Pilar Orero (eds). 2015. *Audio description. New perspectives illustrated*. Amsterdam: John Benjamins.

Mazur, Iwona, and Jan-Louis Kruger (eds). 2012. Pear Stories and Audio Description: Language, perception and cognition across cultures. Special issue *Perspectives. Studies in Translatology*, 20(1).

Matamala, Anna. 2016. The ALST project: technologies for audio description. In *Researching audio description. New approaches*, ed. by Anna Matamala and Pilar Orero, 269-284. London: Palgrave Macmillan.

Matamala, Anna, and Pilar Orero (eds). 2016. *Researching audio description. New approaches*. London: Palgrave Macmillan.

Matamala, Anna. 2019. *Accessibilitat i traducció audiovisual* [Audiovisual translation and accessibility]. Vic: Eumo.

Mazur, Iwona, and Agnieszka Chmiel. 2012. Audio description made to measure: Reflections on interpretation on AD based on the Pear Tree Project Data”. In *Audiovisual translation and media accessibility at the crossroads*, ed. by Aline Remael, Pilar Orero, and Mary Carroll, 173-188. Amsterdam: Rodopi.

Perego, Elisa. 2019. Audio description: evolving recommendations for usable, effective and enjoyable practices. In *The Routledge Handbook of Audiovisual Translation*, ed. By Luis Pérez-González, 114-129. London: Routledge.

Rai, Sonali, Greening, Joan, and Petré Leen. 2010. *A comparative study of audio description guidelines prevalent in different countries*. London: RNIB.

Remael, Aline, Reviere, Nina, and Gert Vercauteren (eds). 2015. *Pictures painted in words. ADLAB audio description guidelines*. Trieste: EUT

Snyder, Joel. 2014. *The visual made verbal. A comprehensive training manual and guide to the history and application of audio description*. Arlington: ACB.