11

Audio description for the screen

María Joaquina Valero Gisbert

1. Introduction: definitions

The concept of “accessibility” applied to architectural barriers (Martínez Calvo, 2003; Orero, 2005) has been widely studied, as well as being the most widespread and well-defined with respect to other problems discussed here. Subsequently, its scope has been extended to the field of information in a broad sense and, in particular, to both the media and their contents (Orero, 2005). In this context, audio description (AD) has developed as an activity for a part of society with problems that mainly affect sight, in order to give those people access to cultural content. AD is proposed as a modality of Audiovisual Translation (AVT) (Gambier, 2004; Díaz Cintas, 2005; Kruger & Orero, 2010), while Mayoral Asensio (2005) in turn considers AVT a specialised modality of Translation Studies.

Other authors confirm this double adhesion of AD to the field of Translation Studies and more specifically to AVT, as acknowledged by Matamala and Orero (2017: 7): “audio description was adopted as a modality in the field of Translation Studies (Orero, 2005) and more accurately in Audiovisual Translation”. Recently, other authors (Gambier & Ramos Pinto, 2018; Greco & Jankowska, 2020) have confirmed this inclusion and its consolidation in Translation Studies. On the other hand, authors such as Yuste Frias (2011b) and Hernández and Mendiluce Cabrera (2009: 1) consider AD as a type of translation “for specific purposes or for groups with special needs”, while other researchers (Mayoral et al., 1988; Mayoral Asensio, 1997) reflect on this translation activity as one in which, in addition to the linguistic code, other systems of communication (iconic, acoustic, proxemic, etc.) complicate the translator’s work. These authors recover the notion of “constrained translation” that had been presented by Titford in 1982 in reference to subtitling and now extended to AVT in general. From our perspective, the audio description modality for the screen as a possible way of providing accessibility for a specific sector of the population, might initially have been considered a type of “constrained translation”, because of the spatial-temporal restrictions to which it is subjected, along with the various codes that combine to make film and other video products understandable. The audiodescribed text provides access to the screen content by synchronising it with the audio track in the spaces selected for it and as one more code among those present in the screen product. However, the technology available to the translator today makes it possible to...
understand translation for the screen in a broad sense as explained earlier. In this context, we agree with the criticism that the denomination “constrained translation” has raised for some time in different forums (Gambier, 2001, 2004; Zabalbeascoa, 2005; Yuste Frías, 2011b), and discard this designation. In fact, the latter author confirms (2011b) that this consideration ignores the role of image in communication and assumes that the translator operates only on the textual dimension.

In our digitalised world, images occupy a privileged place, typically projected through some form of screen. From domestic videophones to the screens used for our leisure or screens on our computers or mobile phones to smartwatches; there are advertising screens in malls and shops in general, in the streets and public places like airports, train or subway stations, city council offices and so on. Displays on cash machines and, at a higher level such as the art world, artists who include screens in their works (installation art), or screens in theatrical performances and, of course, in filmic products.

2. Historical perspectives

As regards AD and its denomination, the term “audio description” seems to have been the most widely used. Others were coined variously as “audio captioning”, “descriptive narration”, “audio vision” and “video description”, though these had fewer adherents, as Navarrete Moreno (2005) and Snyder (2005) explained. More poetically, Snyder defined AD as an art, a kind of “poetry”.

As regards AD’s origins, there are various theories, but the most accredited is the one that places it in the United States in the 1970s (Snyder, 2005). From the professional point of view, this activity was first practiced in the theatrical field (ITC, 2000; Hernández & Mendiluce Cabrera, 2009). Later, in the mid-80s, it arrived in the United Kingdom, where after its beginnings in the theatre, it began to be applied mainly to products that were shown on the small and big screen. There was a shift from theatre to cinema and television. On the other hand, other studies (Orero et al., 2007) place the first practice of this activity in Spain in the middle of the dictator Franco’s regime (for a detailed explanation of the modality used for the audio description, see Orero et al., 2007). These were activities applied to the cinema and carried out live, a practice that was not theorised but which undoubtedly represented an initial stage for its subsequent development. From the point of view of research, as Perego (2014: 15) highlights, the 1975 study by the American Gregory Frazier represents the first academic evidence on the subject.

One aspect to which Perego (2014) also draws attention and which has certainly encouraged the dissemination of and interest in AD, is its role as a support for social and cultural integration. In Europe, various systems have been developed and have made this objective possible. In the Journal on Blindness and Visual Impairment, Navarrete Moreno (1997: 71) lists “the English AUDETEL, the French AUDIOVISION and the Spanish AUDESC promoted by the Spanish National Organisation for the Blind (ONCE) for the last four years”.

Navarrete Moreno (1997) explains the objective of the AUDESC system, namely to give blind people accessibility to image and, in the title of his intervention, he eloquently summarises this operation as “the art of speaking in images”. In their application to the screen, British standards (ITC, 2000: 3) specify the service as follows “a commentary provides a carefully crafted description of actions, locations, body language and facial expressions and is reproduced in the gaps between the normal programme dialogue”. Years later, in Spain, requests from various population groups and also some publications illustrated the situation in that country. After an increasingly evident process of awareness on the part of groups with sensory disabilities,
the activity of AD was standardised with the publication of the Spanish standard UNE 153020 (Aenor, 2005).

Research on AD has become more intense since 2005, especially following its recognition within the framework of Translation Studies. If at the beginning, studies concentrated on the perspective of its application to cinema or television, today we have publications that focus on museums, exhibitions, live shows such as opera and theatre, architecture and natural spaces (see in this regard the rich bibliography presented in Perego, 2014).

Regarding standardisation, the UK published the ITC (2000: 1) which presented “guidelines only, with no absolute rules”. These are indications based on studies carried out between 1991 and 1995 with the aim “to promote the understanding and enjoyment of programmes by sensory impaired people”. A few years later, the well-known Spanish standard, previously mentioned, was published in Spain, describing this practice in very general terms. Similarly, in the rest of Europe official regulations were published to standardise this practice except, for example, in Italy, where each professional body prepares its own rules based, in general, on the 2000 ITC which, as Arma (2014) explains, made an important contribution to the professionals in this sector.

With regard to its practice in the various countries, given the interest now shown by national authorities and also by the film industry, the need to share local experiences in order to arrive at a common, European regulation is becoming more pressing. In this sense, the contributions of various countries to this practice can converge in this direction, as Maszewrowska, Matamala and Orero point out (2014: 3) “various countries in Europe have different strengths and competences. . . . It is thus recommended that European countries exchange experience and share their best practices to shape audio description practice in Europe and bridge the gap between individual countries”, a goal towards which work was done in European projects such as the ADLAB project at Trieste University, whose interesting results Pictures painted in words (2014) can also be consulted on the web (www.adlabproject.eu).

It is clear that the level of specialisation with regard to particular problems is most evident in those countries where practice and research had been carried out earlier and more fruitfully. Orero (2007) points out that since the beginning of the 21st century national and international proposals to standardise professional training in audiovisual accessibility have been advanced, with the European Union dealing to a large extent with the standardisation of communication-related issues. In this sense, the ITC (2000: 5) referred to a possible start of this project when “in 1991 the ITC founded the Audetel consortium which worked with partial financial support from the European Commission to explore all the issues associated with beginning regular broadcasts of described programmes in Europe” (An overview of the different official standards for AD is provided by Perego, 2014: 33). During this first part of the 21st century, research and publications have moved in this direction, for example, Vercauteren (2007) or the more recent work mentioned earlier by Remael et al. (2015) which is, as the authors explain, the result of a three-year (2011–2014) research project on audio description (AD) for the blind and visually impaired, financed by the European Union under the Lifelong Learning Programme (LLP). The basic motivation for the launching of the project was the need to define and create a series of reliable and consistent, research-based guidelines for making arts and media products accessible to the blind and visually impaired through the provision of AD.

(Remael et al., 2015)

This publication is the outcome of the study and collaboration of a number of researchers and professionals from different European countries and universities.
3. Current research and research methods

Research on AD, which has been previously explained, has been widespread in recent years and has also developed in other areas that are not exclusive to its application to the screen, as can be seen in the various chapters included in this volume. Also with regard to our area, that is, the big and small screen (cinema, television, video), studies have been carried out regarding diverse audiovisual categories such as reports, documentaries, television series, sitcoms, soap operas and so on. The characteristic feature of these genres is that they are constructed by combining different systems of meaning. Visual and acoustic elements accompany the linguistic code and together they create a global meaning. Technological development over the years has contributed to a constantly increasing production of these multimodal texts. Multimodality has established itself as a discrete discipline over recent years (see Seizov & Wildfeuer, 2017, and Wildfeuer et al., 2019) and much research has been focused on how to analyse the way in which the different semiotic modes combined convey meaning.

As previously discussed, the multimodal text is the result of the mix of at least two semiotic systems. The different semiotic modes (acoustic and visual) present in this kind of text participate in the creation of the global sense. The translator/describer cannot be unaware of the nature of the various codes that make up the material on which he/she must work. In this regard, Taylor’s study of multimodal transcription (2016) is extremely interesting, as it shows how the analysis of a multimodal text can be carried out by breaking down the text into frames. In these frames the different semiotic modalities that make up the original text are identified and analysed. Although he applies it to the modality of subtitling, there is no doubt about its usefulness for AD.

Following this reflection, we must also consider that depending on the medium in which we express ourselves, verbally or visually, a number of aspects affect meaning, as Kress and van Leeuwen (1996: 2) recognise “what is expressed in language through the choice between different word classes and clause structures, may, in visual communication, be expressed through the choice between different uses of colour or different compositional structures”.

Since almost the beginning of the 21st century, various research methodologies have been followed, some of which derive from a descriptive methodological approach, such as that presented by Jiménez Hurtado and Soler Gallego (2013), combining studies focused on the process of AD with research on its reception. Other works also include contrastive studies; some others involve experimental analysis. It is clear that this object of study can be seen from a multidisciplinary perspective that covers fields ranging from the obvious studies in translation, semiology, film studies, cultural studies and cognitive psychology to theoretical models that are based on narratology (Kruger, 2009; Jiménez Hurtado & Siebel, 2010 and Vercauteren in this volume).

Many of these investigations have focused their attention on various aspects (shots, costume, sounds, light) of AD applied to the screen, such as the description of characters in film, since, as Matamala (2018: 198) explains, they represent an “essential part of a film narrative” and, more specifically, other works have concentrated, for example, on facial expressions (Mazur, 2014), the use of sound (Remael, 2012) and so on.

Other studies combine training and research, as is the case with Matamala (2018), whose work is based on a mixed practical modality. A comparative analysis is made using different versions produced by professionals in three languages (English, Spanish and Catalan) and those produced by students in training (in Spanish and Catalan) providing useful information on the importance of practical experience.
4. Practice

Firstly, riding on the back of research findings and practical experience, the actual production of an AD needs to be analysed. At a basic level the AD of a product transmitted via a screen foresees a series of phases or stages that we can summarise as follows (for further information see Remael et al., 2015 and www.adlabproject.eu):

1) identification of the type of screen product: film, TV series, etc;
2) documentation of process about the author, product and target audience. This phase could be called the hyper-structure or macro-structure of the original or source material (SM);
3) viewing of the product and semiotic analysis of the SM. Taking notes in a glossary about the different codes that make up this material;
4) taking into account the spatial restrictions (timecode), identify the space available for entering visual, non-verbal and acoustic information if required. In particular, the dialogues and soundtrack should not be covered;
5) preparation and writing of the AD;
6) revision of the AD script with regard to cohesion and coherence;
7) viewing of the source material with the AD script. Revision of the whole with a visually impaired person who is familiar with or trained for this work;
8) rehearsal through the voiceover of the AD script with a speaker (human or mechanical) and final control;
9) recording and mixing with the original soundtrack.

With this plan of action in mind, we now refer again to the question of how to deal with the conversion of various semiotic modalities (image, sounds, etc.) into words.

The difficulties faced by the audio describer in conveying sound and images are evident. Kress and van Leeuwen rightly explain (1996: IX): “we wanted to analyse the whole of the texts in which these semiotic modes play a vital role rather than just the verbal part, but also to understand language better”. In this regard, we propose as an example an excerpt from the film Broken Embraces (Almodóvar, 2009) where there is no verbal dialogue between the characters. This is the sequence in which Lena decides to leave Ernesto. As Vilches explains (1984: 73) the moving image is built up by different codes (images, written manifestations) and sub-codes (colours, fonts, etc.). In this regard, it is not superfluous to recall the words of Yuste Frias (2011a) when he emphasises that referring to the chromaticity of images in AD is not only a linguistic question, but has a cultural meaning that depends on the diatopic dimension in which they are manifested.

The meaning of this extract is built up through the iconic and acoustic codes – pictorial work, sound of footsteps, image of Lena and Ernesto’s characters. As far as the former is concerned, the sequence opens with a detailed shot of Motherwell’s painting Je t’aime from where, as Poyato Sánchez (2013) explains:

A descending panoramic view begins until discovering Lena’s face, the “t” of the sentence written in large symbols that presides over the painting hanging, as if it were a dagger, over Lena’s head. This remarkable inscription, where the film seems to warn that the “I” that loves is also a throwing weapon that can hit the “you” loved; an inscription that, as you will see, does not take long to materialise.1

(Poyato Sánchez, 2013: 273)
In addition, the writing contrasts chromatically with other chromatic elements present in
the painting. The acoustic code, on the other hand, is created by the sound of Ernesto’s and
Lena’s steps through close-up detail, steps that come and go. The audio description of this
sequence could be approached from three different perspectives, according to Remael et al.
(2015). The first, as an iconic description, will give the viewer detailed information about the
facial expression so that he/she can mentally create what is happening; the second, the cogni-
tive description, will introduce the interpretative aspect of facial expressions to facilitate the

Figure 11.1  Picture of Motherwell

Figure 11.2  Lena

Figure 11.3  Close-up shot Ernesto’s footsteps
transfer of feelings and their understanding; the third, narrative description, as in the previous one, includes interpretation and introduces, instead of a description, a narrative.

The Spanish and Italian audio descriptions (in English translation) describe the sequence as follows:

---

**AD written by Navarra Decine SL, 2009.**
The name of the audio-describer is not provided.
1:02:47 In the living room. Lena comes out and goes up the stairs to the upstairs bedroom. 1:02:54 She sits on the edge of the bed. Waiting. 1:03:00 Ernesto senior reaches the bedroom door. He stops. 1:03:10 Lena is tense inside. 1:03:18 But Ernesto Martel turns around and leaves. 1:03:21 Lena hears him walking away. 1:03:23 Lena is wearing a fancy red passion dress with high heels. 1:03:26 She walks down the corridor. 1:03:37 Lena crosses in front of Ernesto's office. She watches him. 1:03:45 Ernesto is sitting in an armchair with a glass of whiskey in his hands. He is having a drink. 1:03:50 They look at each other. Lena crosses in front of him with a suitcase in her hand. 1:04:09 He goes to the stairs.²

---

**AD written by Cinema senza Barriere – Milano.** The name of the audio-describer is not provided, nor the year of production.
1 TCR 04 10 00 22 Lena leaves the living room, goes up the stairs to the upstairs; she has a red tight-fitting dress. Then we see her sitting at the table. The camera frames Ernesto's feet, then her feet, red high-heeled shoes with no stockings; in the darkness of his study he is drinking something. . .³

---

Both audio-written scripts contain important differences: the number of words, the use of film language, the use of perception verbs related to the act of looking, the explanation of sounds “Lena hears him walking away” or the lack of information about the visual composition of the sequence. As for the pictorial work, the black colour of the written text *Je t’aime* contrasts with the warm background of the work, an expressive antithesis between the words and the colours used. In both descriptions, what Kress and van Leeuwen (1996: 19) point out is that the concept “‘subjective’ and ‘objective’ can be applied. . .: they belong to the meaning.
potential of a culture and its society”. On the other hand, as Vilches (1984) states regarding the image, it will depend on the reader’s visual and verbal skills, on his or her ability to interpret the images given by the author of the text. And it also means taking into account, as Kress and van Leeuwen explain, that the language, whether in speech or writing, has always existed as just one mode in the ensemble of modes involved in the production of texts, spoken or written. A spoken text is never just verbal, but also visual, combining with modes such as facial expression, gesture, posture and other forms of self-presentation. A written text, similarly, involves more than language: it is written on something, on some material (paper, wood, vellum, stone, metal, rock, etc.) and it is written with something (gold, ink, (en)gravings, dots of paint, etc.); with letters formed as types of font, influenced by aesthetic, psychological, pragmatic and other considerations and with layout imposed on the material substance, whether on the page, the computer screen or a polished brass plaque. Yet the multimodality of written texts has, by and large, been ignored.

(Kress & van Leeuwen, 1996: 41)

Considering the different “ways of meaning” present in the audiovisual text, we cannot forget that the understanding of this multimodal text by the audio-describer or reader in the Vilchian sense (1984) is a pragmatic task. The reader or viewer updates the visual text according to their skills and their own knowledge. However, this does not mean that one can make unlimited readings because, as Vilches explains (1984), reading a text is done on the basis of common reading behaviour rules for a given community. Among the reader’s skills, perceptive-syntactic and semantic, we highlight the modal competence which is based on an authorial intentionality and the ability to carry it out. The importance (of modal competence) lies in what the author wants to show and the view of the reader who will actualize or reveal it. In this sense, the methodology suggested by Taylor (2003: 191) can be very useful, as we have already pointed out. It consists in “breaking down a film into single frames/shots/phases and analysing all the semiotic modalities operating in each frame/shot/phase” (see Valero Gisbert, 2021).

Interestingly, the study of AD can be linked to the discipline of Cultural Studies (Hall et al., 1980; García Canclini, 1995; Bennett, 1997), since they can provide the necessary knowledge to understand or, as Herrero & Escobar explain:

“reading” the filmic text and rendering the cultural messages. The analysis of cinematography, mise-en-scène, editing, sound, genre conventions and narrative construction provides a solid ground to examine the way in which social, cultural, political and historical representations are conveyed in films and how they are intertwined.

(Herrero & Escobar, 2018: 52)

This consideration leads us to propose an approach to AD from the perspective of Social Semiotics, where meaning is constructed thanks to social rules which, in turn, can be modified through social interaction. From the point of view of AD – consisting of the transmission of the iconic content especially, but not exclusively, through the linguistic code – all the codes involved in the creation of meaning in a film are, as explained by Jewit and Henriksen (2016: 150), “connected to the social context of use”. This is very important from the perspective of the transfer of cultural content through AD. Various researchers discuss this aspect and point out specific solutions (Matamala & Rami, 2009; Sanderson, 2011; Sanz-Moreno, 2017, 2018; Valero Gisbert, 2020).
Finally, Hernández and Mendiluce Cabrera (2009) make a proposal that goes beyond the visually impaired and focuses on an audience with acquired cognitive impairment, which affects mainly older people, but also children, with a view to creating a better understanding of the information they receive.

The same can be true of children or elderly people, who may find some of those audio-visual comments useful for drawing their attention to meaningful facts about the film which otherwise might have gone unnoticed, or for teaching or reminding them of relevant aspects of culture, art, literature or society.

(Hernández & Mendiluce Cabrera, 2009: 3)

In line with these works, other studies have identified a type of public to which AD could help as cognitive support, namely immigrants or illiterate people and, again, children. It is a tool for understanding perhaps new realities and acquiring vocabulary as Vermeulen and Ibáñez Moreno (2013) show. Remael et al. (2015: 15) already point out that “AD is being used by an increasingly large group of sighted viewers for an equally varied number of reasons. Immigrants may use AD to learn the language of their host country, children may use AD as they are acquiring languages”.

On the other hand, studies that focus on the target group are becoming increasingly visible. In a reception study conducted by Fryer and Freeman (2012a) on the use of filmic language, it is concluded that, except for the innate blind, other visually impaired participants responded by validating this practice, despite the recommendations contained in the standards for AD.

Considering what has been illustrated thus far and the four techniques that Perego (2014) identifies for the analysis of a film (denotative, expressive, symbolic and aesthetic), we wonder if an audio introduction that provides information about these aspects for a better enjoyment of its content could be appropriate for the audience with innate or acquired visual disability. It is clear that we are entering the field of future research work where Reception Studies can add important approaches to future experimentation. The need to encourage this kind of receptor-based research is becoming increasingly evident. We point out Braun’s study (2008) on the amount of information or Sanz-Moreno’s (2018) on the efficiency in the transmission of cultural contents.

Reception Studies exist, although they have not been very frequent, partly because of the evident difficulties (Orero et al., 2018) in organising collaboration according to age, sex, training and so on. The Jekat, Prontera and Bale (2015) trial used a methodology for a reception study in which the information conveyed by two kinds of AD (descriptive and interpretative) is compared through two groups of recipients with very different visual characteristics, that is, visually impaired people and others without sight problems. The results show that users have the same perception, and that interpretative AD is preferred over descriptive AD because it facilitates or improves the transmission of information. These are results that go in the direction of further studies (Walczak & Fryer, 2017). Following these considerations, we move on to approach future or developing lines of research.

5. Future directions

Matamala and Orero (2017) in “Audio Description and Social Acceptability” state that since the origins of AD and especially since it has been recognised as a modality of AVT, research has increased with a substantial number of academic publications. They show (2017: 7) that there has been a change in the approach to this modality from research based on a “superficial
study on a clinical description and classification of the physiological features” to works that adopt the perspective of citizens’ rights, for example considering AD as an “artistic cultural product”, allowing them to study it from a sociological perspective.

Some studies, as mentioned previously, have focused on identifying potential users different from those with eye-related pathologies. Audio descriptions, made for the cinema or television, addressed to visually impaired people, can be used, as explained by Hernández and Mendiluce Cabrera (2009), for people with innate or acquired cognitive disabilities, although, as reported by Orero (2005), the problem of access to audiovisual communication by this group is still not very visible. Currently, there are no separate audio descriptions for each of these sectors of the population. We hope there will be in the future thanks to new research projects that will establish whether the current standards are sufficient. In fact, a study of the needs and problems presented by each of these social groups would be necessary in order to understand and produce the type of AD that best suits their needs. A person with cognitive impairment may not find the same type of information useful as that received by a person who cannot see. Some studies, such as Bernabé Caro & Orero (2019), forward a proposal that facilitates understanding, and the EASIT project (Easy Access for Social Inclusion Training. http://pagines.uab.cat/easit) is working in this direction. Until now, most studies have focused solely on the needs of a visually impaired recipient. Other possible beneficiaries are people with autism whom AD can help identify emotions (Fellowes, 2012), for example, entering into a domain in which the notion of neurodiversity emerges (Singer, 1998). It would be necessary to focus on each of these groups and propose a diversification of audio descriptions depending on the audiences or people to whom they are addressed.

On the other hand, and given the speed with which society is increasingly consuming “screen products”, normally sighted people, who perform tasks that prevent them from looking at a screen, such as driving, can make use of audio described films. Both in this case and, above all, in the previous ones, reception studies will reveal a fundamental point of reflection on the future of this practice.

The methodologies of studies applied to other modalities of AVT, such as subtitling, can also be extended to AD, as suggested by the authors of the research by Kruger et al. (2017) in which the authors consider new ways, such as that experienced through the tool “electroencephalography” (EEG) to measure the degree of immersion of the public in a story.

Finally, we refer to the need to explore in future research the communication that develops on the screen in the broad sense that we alluded to at the beginning of this chapter. In the field of audio description of video games, research is becoming increasingly visible (Mangiron & Orero, 2012; Maszerowska & Mangiron, 2014). As regards this area, Mangiron and Zhang in this volume. However, there are not many studies that have dealt with AD in the different types of screens that we have mentioned. Each one has its own characteristics and functions (playful, informative, services in general, etc.). The development must take into account the specificity of the environment, so it is necessary to study the intrinsic characteristics of the object in relation to its function (for example, for installation art, we need to recognise the role of the screen in relation to the environment where it is exhibited and, in particular, the involvement of the subject in an immersive and/or interactive experience). For this type of AD, see Perego and Taylor in this volume.

As regards AD on the screens of different devices, such as computers or mobile phones, automatic programmes for reading written text are used (Guerreiro et al., 2017). Regarding the reception of synthetic voices that imitate the human voice (Fernández-Torné & Matamala, 2015), we enter a multidisciplinary field where computer engineering (Cryer & Home, 2008) – to cite one discipline – has much to offer. However, there is a lack of research that incorporates
the description of the images that accompany the written text on this type of screen and, subse-
sequently, that provides information on the fruitfulness of these contents (Walczak and Gallardo 
in this volume).

6. Other directions

Pedagogy applied to AD is one of the lines of research that is becoming increasingly visible. 
Vermeulen and Ibáñez Moreno (2013), through their study, try to answer whether AD can be 
a good tool to increase the lexical and phraseological knowledge of students of an L2 (second 
language). Perego’s study (2014) is placed in this same pedagogical context, although with a 
different purpose since it relates it to cultural content. According to the author, the enjoyment 
of a film in its original version with audio description facilitates the understanding of expres-
sions associated with an image.

With regard to this proposal, we present an example for the transfer of cultural content, 
in this case in the field of gastronomy. This sample comes from the Spanish AD of the film 
Women on the Verge of a Nervous Breakdown (Almodóvar, 1988). This is the sequence in 
which Pepa is preparing a typical dish of Spanish cuisine.

The Spanish audio description reads: “cuts up tomatoes, throws them in a blender to make 
gazpacho”. The explicitness of the dish is undoubtedly an interesting fact to transmit that 
content if we think of an AD created ad hoc for this specific public, that is, students of Spanish 
as a foreign language. However, if the AD is aimed at a visually impaired Spanish audience, 
it should not be made explicit; through the description of Pepa’s action, the visually impaired 
should be able to create the mental image of the dish she is cooking. Taking into account what 
we explained earlier, it would be necessary to propose audio descriptions suitable for users 
with different needs. In this respect, we also suggest here the question of the need to diversify 
ADs according to the public or audience they are addressing.

On the one hand, some researchers bring together the previous proposals and reflect on 
the role of AD in education, specifically the study by Herrero and Escobar (2018: 30) which 
presents a pedagogical model in which AD is used to “help students develop linguistic, cul-
tural and intercultural competences while encouraging the aesthetic appreciation of films 
as cultural objects . . .”. They explain that it is a yet unexplored angle (2018: 39): “films 
are still an under-exploited resource for promoting intercultural competence and developing 
learners’ critical thinking skills”. On the other hand, it can be especially useful when 
we have to face products from culturally distant worlds, such as the East and the West. 
Making some sociocultural uses explicit will help to spread that culture and to understand 
certain behaviours. As Kress and van Leeuwen point out (1996: 4), “Visual language is 
not – despite assumptions to the contrary – transparent and universally understood; it is 
culturally specific”. In this regard, orientations in the field of research are proposed by 
Herrador Molina (2006), López Vera (2006) or Jankowska (2015) which explore the pos-
sibilities offered by the translation of AD. This line of study is closely related to Cultural 
Studies, Film Studies and Reception Studies. López Vera, starting from Benecke’s (2004) 
and Hyks (2005) proposals, concludes that in economic terms it can constitute a good 
practice, since it involves a lower cost and requires less time for elaboration. It contrasts 
partly with Hyks view when she explains that the process of adaptation may take longer 
than its creation ex novo. For her part, Jankowska points out, along the lines of López Vera, 
that it can be a good starting point, although she stresses that it is not a question of merely 
translating a written text, since such an operation must also take into account the images 
(Oncins in this volume).
In the field of Reception Studies, it would be interesting to analyse Kruger’s (2010) proposal in which he studies literary narrative by comparing it with film. Kruger proposes the term “audio narration” (AN) as a substitute for the term audio description. This author explains that guidelines usually refer to what can be seen and not to how or why (2010: 237–238), “which, in most guidelines, would constitute a subjective interpretation that should be avoided”. Kruger claims that (2010: 237) a “re-narrativisation will always be based on an interpretation of the visual presentation”.

Another field of study which has not been mentioned and which has important repercussions on the reception of AD is noted by Arma (2014). We refer to the phase in which the speaker records the AD. Arma explains the need for specific research on the kind of information contained in the audio-written script and which is addressed to the speaker with the task of verbalising it orally (e.g., reading quickly, slowly, etc.).

On the other hand, new technologies applied to AD are starting to be explored. These are studies related to Virtual Reality and audio description as Fryer and Freeman (2012b), Fidyka and Matamala (2018) or Orero (this volume on technological advances). In particular, Fryer and Freeman (2012b) propose its use in live shows, where their use would be limited to the beginning of the show. The authors (2012b: 20) suggest the need for proposals of analysis on “how variation in AD style affects presence for those with impaired sight and how it might need to be adapted for virtual reality, perhaps not only through linguistic adaptation but also through a more creative use of sound”. Suggestions have been added to the work of Montagud et al. (2020) with the intention of understanding the needs, challenges and limitations for making VR360 experiences accessible.

To conclude, we would like to note a recent line of research that is taking its first steps and that relates to the creation of audio descriptions with partly semi-automated means. This is the study by Braun and Starr (2019 and this volume) who, in turn, based on innovative and highly technological studies oriented towards producing descriptions of still images through the machine, try to discover whether it can be applied to AD.

7. Conclusion

In this chapter we have dealt with AD applied to the screen. Within the framework of Translation Studies and especially Audiovisual Translation, AD emerges as one of the most interesting modalities in the field of translation in the broadest sense, due to the kind of “text” it has to deal with. Its purpose is related to accessibility in the sense that it establishes a bridge between the object and its target audience. AD makes possible a transmission of information from a multimodal text to people, principally, with visual disabilities. However, as has been seen throughout this chapter, both the group of people who can be addressed has grown in recent years and its applications have been diversified. AD has gone from being a product designed especially for people with visual disabilities to also being used by people without disabilities. It has a role in conveying cultural content in general, teaching languages, as a vehicle for literacy, support for the elderly or for people with different degrees of cognitive problems.

A brief historical evolution has been traced, showing the importance of research into this practice, which is born of a need, the oldest one, to communicate. In this sense, a journey has been made through different research methodologies, works that show the complexity of this activity, in particular due to its interdisciplinary nature, since it covers, as we have seen, humanistic and technological disciplines. This combination is opening up the field of application to other areas, as we have pointed out in the section dedicated to future developments and other surprising directions.
Notes

1 Our translation from the original quote: se inicia una panorámica descendente hasta descubrir el rostro de Lena, la “t” de la frase escrita con grandes caracteres que preside el cuadro pendiendo, como si de una daga se tratase, sobre la cabeza de Lena. Notable inscripción ésta donde el filme parece advertir que ese “yo” que ama es también un arma arrojadiza que puede golpear al “tú” amado; inscripción que, como se verá, no tarda en materializarse.

2 Our translation from the original quote: 1:02:47 En el salón. Lena sale y sube las escaleras hacia el dormitorio del piso superior. 1:02:54 Se sienta en el borde de la cama. Esperando. 01:03:00 Ernesto padre llega hasta la puerta del dormitorio. Se detiene. 1:03:10 Lena está tensa en el interior.1:03:18 Pero Ernesto Martel se da la vuelta y se marcha. 1:03:23 Lena le escucha alejarse. 1:03:26 Recorre el pasillo. 1:03:37 Lena cruza frente al despacho de Ernesto. Le observa. 1:03:45 Ernesto está sentado en un sillón con un vaso de whisky entre las manos. Bebe un trago. 1:03:50 Se miran. Lena cruza ante él con una maleta en la mano. 1:04:09 Se dirige a las escaleras.

3 Our translation from the original quote: 1 TCR 04 10 00 22 Lena esce dal salotto, sale lo scalone di casa. che porta al piano di sopra; ha un tailleur rosso attillato. Poi la vediamo seduta a tavola. La macchina da presa inquadra i piedi di Ernesto, poi quelli di lei, scarpe rosse col tacco alto, senza calze; nel buio del suo studio lui sta bevendo qualcosa.

8. Further reading


See this article for a detailed explanation of the modality used for the audio description.


A rich bibliography and an overview of the different official standards for AD is provided by this volume.


9. References


Yuste Frías, J. (2011b). Traducir para la pantalla: El traductor entre el texto y la imagen. En Di Giovanni, E. (Ed.), *Diálogos intertextuales 5: Between text and receiver: Translation and accessibility*. Entre...

10. Filmography