CADERNOS de TRADUÇÃO

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

Alejandro Romero-Muñoz
Universitat Jaume I
Castellón de la Plana, Spain
alromero@uji.es
https://orcid.org/0000-0002-5869-0652

Abstract: The appropriateness of subjectivity or objectivity in audio description (AD) is an ongoing debate where the professional and the academic perspectives seem to collide. Whereas some professional guidelines support the view that an objective AD is the advisable option, some scholars have pointed out the advantages of more creative alternatives. Meanwhile, part of AD research is making inroads into experimental methodologies: AD and eye-tracking, text-to-speech AD, the automatic generation of AD, and physiological measurements applied to AD, such as heart rate, cortisol levels, or skin conductance response, among others. In this vein, our proposal presents an experimental pilot study using a heart rate sensor to measure partially sighted participants' response to different combinations of subjective and objective multimodal components within the AD script. To do so, objective and subjective multimodal components from an AD clip from Netflix's *Money Heist* were analysed. Afterwards, objective components were turned subjective (and the other way around), which resulted in two versions of the clip in subjective-objective terms: the original and the creative one. Both versions were recorded by a professional studio, participants were randomly assigned one of the clips, and they orally completed a demographic survey, a presence test, and a comprehension test followed by a semi-structured interview. Moreover, the UPTIVO Belt-D heart rate sensor was used to measure beats per minute, average heart rate, maximum heart rate, minimum heart rate, etc. Preliminary results were triangulated with the tests and the interview to obtain empirical data about the combination of subjective and objective multimodal components in AD that induces the highest heart rate and presence levels, enables better comprehension and is best regarded by users. Future research following this methodological design and involving an appropriate number of participants could provide empirical evidence about the immersive combination of objectivity and subjectivity in AD to be considered by guidelines.
Keywords: multimodality; audio description; creativity; heart rate; experimental research.

1. Introduction

Audio description (AD) is an accessible audiovisual translation (AVT) and media accessibility (MA) mode that makes audiovisual products (such as films, series, or documentaries) and events (theater, opera, or museums) accessible by means of a verbal narration that translates visual elements (as well as some aural elements that are difficult to interpret) primarily for visually impaired or partially sighted users. Jankowska (2015) as well as Perego and Pacinotti (2020) state that rudimentary forms of AD can be linked to the explanation of images and inscriptions since ancient times, but AD as a professional discipline appeared during the 1960s in the USA with Chet Avery in the American Department of Education. According to Mazur (2020), AD in the USA is more systematically found in 1970s and 1980s theater, which later expanded to UK theater, and then to television (Spain and Italy) and cinema (Poland and Germany). Perego and Pacinotti (2020) point to Gregory Frazier's *The Autobiography of Miss Jane Pittman* (1975) as the starting point as a discipline, where AD as a concept is first developed, but it will not be until the late 1990s and early 2000s that AD begins to receive more scholarly attention (Greco & Jankowska, 2020). One of the crucial aspects that has been tackled both professionally and scholarly is the issue of objectivity in AD, so in the following section we will discuss professional and academic views on the topic. After that, we will develop the methodological design and preliminary results of a pilot heart rate experiment to study the most immersive combination of objective and subjective multimodal components in AD that are best received by end-users.

2. Objectivity in audio description. Profession versus research.

Very diverse AD documents, usually called "guidelines", "quality standards", or "norms", regulate the contents of the AD script to different extents, which helps audio describers in their decision making. The concept of "norm" is related to the notion of "standard" and refers to the prescriptive instructions about how to audio describe in certain circumstances. According to Iturregui-Gallardo (2019), this type of document is created by a national or international official standardization agency of a professional nature that intends to homogenize AD. The Spanish *UNE 153020* norm (AENOR, 2005) is an example of a national standard, since it was created by the Agencia Española de Normalización y Certificación (AENOR) and developed by a commission of ONCE members (the national organization for the blind and visually impaired), audio describers drawing on their professional experience, and surveys. In 2000 in the United Kingdom, Ofcom published a revised version of the 1996 *ITC Guidance on Standards for Audio Description*, a document that originates from the British reception study Audio Description on Television (AUDETEL) project that took place between 1992 and 1995 (Jankowska, 2015). In 2021, Ofcom published the *Guidelines on the Provision of Television Access Services* (updated in 2024), which covers AD, subtitling for the deaf and hard of hearing, and sign language interpreting. In the United States, two national standards can be mentioned: the *Audio Description Standards*, created by the American Council of the Blind (ACB) in 2009, and the *Standards for Audio Description and Code of Professional Conduct for Describers* created in 2009 by the Audio Description Coalition. From an international perspective, we could highlight the *ISO/IEC 200071-21* standard (2015), a document created by the International Organization for

Standardization (ISO) to address the creation of AD and its contents. Another international standard is the *European Blind Union (EBU) Handbook for High Quality Audio Description on Screen* (Hyks & Blokland, 2023), which addresses issues such as the use of verb tenses, articles, the vocabulary that should be used, character naming, their physical appearance, the issue of diversity, the levels of importance of certain activities (principal, secondary, or non-essential), content selection, and so on. Similarly, we should mention the notion of "guidelines", which are "recommendations produced by private and public companies, professionals or scholars" (Iturregui-Gallardo, 2019, p. 42). Guidelines can be created by private companies (for instance, Netflix), by associations or institutions (such as the American Council of the Blind), or they can have an academic nature, like the European project ADLAB, which results from the European reception studies on media accessibility Digital Television for All (DTV4ALL) and Hybrid Broadcast Broadband for All (HBB4ALL).

Despite their diversity, all of them exert a certain sense of prescription over AD. In this sense, Pedersen (2020, p. 426) talks about experience-based norms, in which "practitioners […] reported on their experience of their work", and where, drawing on introspection and reflection, fundamental good practices for a discipline are described, which resonates with some of the national standards already explained. After that, Pedersen (2020) includes descriptive norms, which result from empirical research based on case studies or corpus studies (such as ADLAB). Finally, Pedersen (2020) identifies the experimental norms, which confirm or question the contents of previous norms. In our view, while current guidelines and standards have an evident benefit for the homogenization of AD, experimental studies might help AD to better adapt their contents to end-users' needs. This urge for experimental studies goes in consonance with the latest developments in AD research. In the 1990s and 2000s, early studies revolved around AD in different countries, their linguistic and textual traits, and their standards or guidelines, which gave rise to studies on alternatives to AD in different aspects, such as objectivity (Perego & Pacinotti, 2020). Progressively, research on AD has expanded into three main categories (Mazur, 2020): text-based, reception, and experimental studies. Given the nature of this paper, it is worth focusing on the myriad of AD research from an experimental perspective: AD and eye-tracking (Krejtz et al., 2012; Kruger, 2012; Orero & Vilaró, 2012; Di Giovanni, 2014; Mazur & Chmiel, 2016; Krejtz et al., 2024), text-to-speech AD (Szarkowska, 2011), the automatic generation of AD (Delgado et al., 2015; Hasegawa-Johnson et al., 2017), and physiological measurements applied to AD, such as heart-rate (Jankowska et al., 2022), cortisol levels (Ramos Caro, 2015, 2016; Rojo López et al., 2021), or skin conductance response (Matamala et al., 2020), among others.

Taking into account the number of documents that regulate AD from a professional perspective and the flourishing of research on this MA mode, it seems convenient to tackle the widely discussed notion of objectivity and the creative alternatives to it. We can find recurrent allusions to objectivity in many guidelines, it is a constant topic at AD congresses and in special issues from international journals, and many papers reflect on whether AD should opt for objective or subjective renditions. However, the disagreement revolves around two opposing views: professional AD guidelines have traditionally preferred objective descriptions (Romero-Fresco & Chaume, 2022), whereas some AVT and MA scholars prefer more creative alternatives.

Regarding the professional perspective on objectivity, national AD standards are a clear example of this viewpoint. The Spanish *UNE 153020* norm advises not to describe any subjective

point of view (AENOR, 2005). The French *La Charte de l'audiodescription* suggests that the description must be done in an objective way and not to use subjective adjectives (Morisset & Gonant, 2008). In the UK, the *ITC Guidance* (2000) states that AD should not reflect the personal view of the describer, whereas *Ofcom's Guidelines on the Provision of Television Access Services* (2024, p. 8) state that "adverbs are a useful shorthand to describing emotions and actions but should not be subjective". In the US, a similar recommendation can be found in the *Standards for Audio Description and Code of Professional Conduct for Describers* (Audio Description Coalition, 2009, p. 2), which suggest to "describe objectively", as well as in the American *Audio Description Standards* (California Audio Describers Alliance, 2009, p. 9) when they point out that "the best audio describers objectively recount the visual aspects of an image". In Brazil, the *Guia para Produções Audiovisuais Acessíveis* (Naves et al., 2016) states that AD language must be objective, simple, and succinct (yet vivid and imaginative), while the *Acessibilidade na comunicação: Audiodescrição* (ABNT, 2018) suggests that the AD script should be coherent, cohesive, fluid, with an objective syntax, etc. Moreover, Joel Snyder's (2014, p. 163) *The Visual Made Verbal* asserts that "the best audio describer is sometimes referred to as a *verbal camera lens*, objectively recounting visual aspects of an event". Finally, Netflix's (2024, p. 1–2) guidelines state that "description should be factual" and "description should not be opinionated".

As for the academic perspective, there have been considerable attempts to explore more creative alternatives that deviate from objectivity, which aligns with the concept of "creative media accessibility", which is "those practices that not only attempt to provide access for the users of a film or a play, but also seek to become an artistic contribution in their own right and to enhance user experience in a creative or imaginative way" (Romero-Fresco & Chaume, 2022, p. 84). Even if these proposals have been labelled as "interpretative", "creative", "narrative", or "subjective", we believe that the term "creative" encompasses all of them as examples of creative media accessibility. In this light, according to Jankowska (2015, p. 22), "a hundred per cent objectivity is impossible, and AD is always subjective, since it is a choice made by a particular audio describer". ADLAB states that "AD too is always subjective to some extent since it is based on the interpretation of the audio describer" (Remael et al., 2014, p. 8). Mazur (2020, p. 235) explains that "although AD guidelines generally advise against using subjective judgment in description, in some cases an objective description may be difficult". We believe that the intersemiotic transfer in AD entails some textual subjectivity related to what to include in the script, which differs from a personal subjectivity related to how to describe those elements. This difference appears in ADLAB (2013), where a distinction is made between personal interpretation (subjectivity) and text-based interpretation (objectivity). In our view, ADLAB's text-based interpretation should not be equated to objectivity, but it is rather proof that complete objectivity is never possible in AD.

Focusing on the creative alternatives proposed by scholars, Kruger (2010) introduced the notion of the "descriptive-narrative continuum" with three types of AD: an explicitly descriptive AD, an AD with some narrative markers and subjective interpretation, and finally an audio narration that favors a coherent narration. Szarkowska (2013) proposed an *auteur* description that includes vivid and emotional language, additional information about the characters, their emotions, as well as actions and settings not necessarily visible on screen. Jankowska (2015) suggests a creative audio description that includes vivid language, metaphors, emotions, and film language. Walczak (2017)

also proposes a creative description, which uses *mise-en-shot* elements as well as intensified, vivid, and emotional vocabulary. Moreover, a traditional notion that can be found in research on creative AD is the concept of "style". In this sense, Fryer and Freeman (2012) propose a "cinematic AD" style that includes filmic language. Bardini (2020) distinguishes conventional, cinematographic, and narrative AD styles. Holsanova (2016) studies a descriptive AD style that focuses on spatial and visual details, and a narrative style that covers temporal and dynamic elements. Soler Gallego and Luque Colmenero (2023) study AD for museums and compare a standard style to minority styles, such as what they call "gist style". Finally, Rizzo and Spinzi (2023) study creativity as a multidimensional tool in AD, and to do so they investigated whether and how creativity is used to transmit culture-specific references in "authorial AD".

In conclusion, there seem to be two points of view regarding objectivity in AD. Firstly, we can identify a professional perspective present in most guidelines where objectivity is a desirable option. On the other hand, an academic perspective rejects the idea of complete objectivity being possible and encourages more creative alternatives (where "creative" is considered to be a deviation from objectivity). Moreover, part of AD research seems to be moving towards experimental methodologies, so in this paper we will contribute to the objectivity debate following some of the latest trends in AD research by presenting an experimental proposal using a heart rate sensor to measure partially sighted participants' response to different combinations of subjective and objective multimodal components, as will be fully explained in the following section.

3. Methodology

This proposal draws on previous descriptive research (Romero-Muñoz, 2023, 2025) where the multimodal components (i.e., the multimodal configuration) of some filmic AD scripts from some streaming platforms proved to revolve around four sources of information: movement, iconography, spatial-temporal changes, and textual information. Following the recent trends in AD research, an experimental reception study seemed an adequate option to properly tackle end-users' needs beyond prescriptive or descriptive views. More specifically, our proposal is a heart rate study among partially sighted participants that seeks to determine the most immersive combination of objective and subjective multimodal elements triggering the highest levels of presence, comprehension and heart rate measurements.

Even though reception studies are essential to provide AVT researchers with information about users' preferences, which can help determine what processes best fit certain audiovisual products (Mangiron, 2022), they are not very frequent in our tradition (Orrego-Carmona, 2019). In reception studies, "different research methods have been used, ranging from surveys and questionnaires to determine users' preferences, to more experimental studies using eye-tracking and other biometric indicators, such as heart rate" (Mangiron, 2022, p. 416). As Orrego-Carmona (2019) explains, the main research methods in reception studies applied to AVT are questionnaires, biometrical indicators (such as eye-tracking), interviews, direct observation and focal groups. It must be clarified that these methods are usually mixed, so in this case we propose a combination of questionnaires (quantitative data), a heart rate sensor (quantitative data), and a semi-structed interview (qualitative data), as will be developed in the following subsections.

3.1 Questionnaires and interview

Questionnaires are a method than can be found in some AVT research since the emergence of the sociological turn in Translation Studies, although they are also present in experimental research (Saldanha & O'Brien, 2013). Although similar (yet slightly different) concepts coexist, such as "survey", "test", and "questionnaire", we prefer the term "questionnaire" here, which is a list of questions in a format that allows structuring general information, opinions, attitudes, behaviors, and other data about participants in a standardized way, thus allowing this data to be generalized to a larger population (Saldanha & O'Brien, 2013). Given their subjective nature, questionnaires in experimental studies are normally used before or after the stimulus, while physiological measurements (such as heartbeat) are used during the stimulus (Iturregui-Gallardo, 2019). Questionnaires should be short to ensure completion (Saldanha & O'Brien, 2013) and they need to be specifically designed to obtain the information needed. According to Mazur and Chmiel (2016, p. 102), "comprehension questions may bring more objective data than preference questions since the former test comprehension and not personal likes and dislikes". In this light, three questionnaires were used in this experiment: a demographic questionnaire (before the stimulus), a comprehension questionnaire (after the stimulus), and a presence questionnaire (after the stimulus). Since participants were partially sighted, questionnaires were read out loud by the researcher and the answers were recorded. Demographic questionnaires help researchers specify the participants' profile, so this experiment opted for questions about their gender, birth date, academic background, their type of vision loss (total, partial, congenital, or acquired), how often they consumed audiovisual products, whether and how they used AD, etc. Regarding the comprehension questionnaire, it included both multiple-choice and true or false questions, and participants answered it right after the stimulus so that they did not forget about the fragment (a clip from *Money Heist*). Finally, in this experiment, presence is an essential notion to measure participants' immersion, so the ITC Sense of Presence Inventory or ITC-SOPI (Lessiter et al., 2001) was selected given its frequency in other AVT and MA studies. More specifically, we used the ITC-SOPI short form used by Fryer and Freeman (2014), which includes ten questions about four dimensions (physical space, engagement, ecological validity, and negative effects) that must be answered using one-to-five Likert scales.

However, questionnaires may not be the best instrument to address explanatory data, such as experiences or opinions (Saldanha & O'Brien, 2013), so the experiment design also resorted to qualitative data retrieved from interviews, since they allow access to participants' thoughts and opinions about a specific topic (Saldanha & O'Brien, 2013). In this experiment, we used a semi-structured interview where participants were asked whether they were familiar with the fragment, whether they could provide a summary of the contents, whether they remembered any reference to multimodal contents (movements, iconography, spatial-temporal changes or text appearing on screen), whether they found the fragment immersive, their opinion about the AD voice, whether they considered that the AD addressed their needs, whether they would change anything about the AD, and finally whether they preferred objective or creative versions of AD. While these questions were prepared beforehand, some changes were allowed depending on the participant's interaction. The interview was recorded and processed afterwards with the qualitative analysis software ATLAS.ti.

3.2 Heart rate and immersion

In experimental reception studies, it is frequent to complement questionnaire data with physiological measurements (Orrego-Carmona, 2019). In this case, we complemented questionnaires and the interview data with heart rate measurements to study the notion of immersion in AD. More specifically, the aim is to find the most immersive combination of subjective and objective multimodal components obtaining the highest results with the heart rate sensor, questionnaires, and the interview. In this way, we seek to come up with the appropriate formula so that AD can be not only a source of information, but also a source of entertainment and enjoyment. This methodological combination follows the trend suggested by Kruger and Doherty (2018, p. 91), according to whom, in AD, "a clear trend is also emerging to investigate subjective audience reception and immersion as a result of different styles of AD, as well as to use eye tracking of sighted viewers to determine areas of salience on screen that could be used in deciding what to audio describe".

As Mazur and Chmiel (2016) state, there have been some reception studies exploring users' acceptance of new AD solutions: text-to-speech AD (Szarkowska & Mączyńska, 2011; Szarkowska & Jankowska, 2012), *auteur* AD (Szarkowska & Wasylczyk, 2014), AD in verse (Udo & Fels, 2009), AD narrated in first person (Fels et al., 2006), the congruence of AD voices with the filmic content (Iglesias Fernández et al., 2015), intonation (Cabeza-Caceres, 2013), etc. However, not all of these reception studies are experimental. As Díaz Cintas and Szarkowska (2020) remind us, some AVT and MA researchers have focused on ways to better understand audiences' attitudes and behaviors towards certain AVT modes by means of offline methodologies, such as questionnaires, interviews, focal groups, or direct observation. However, the cognitive turn has facilitated research focused on cognitive processes, which, in turn, has led to the use of online and offline experimental methodologies, among which González and Jankowska (2024) highlight the use of physiological measurements, such as eye-tracking, electroencephalography (EEG), galvanic skin response, and heart rate, to study issues such as information processing, reading behavior, attention location, cognitive load, comprehension, and emotions. In our case, we will use physiological measurements stemming from heart rate data in immersive AD.

We must state that the concept of "immersion" can be elusive, since some researchers often resort to notions like "presence", "enjoyment", "entertainment", "engagement", etc. According to Walczak (2017, p. 26), "immersion is a term used to describe the sensation of the audience being plunged into the story world, experiencing the mediated environment as if it was unmediated". Cavallo and Fryer (2022, p. 132) explain that immersion is "an objective assessment relating to the degree to which the individual is cut off from the real world such that their sensory inputs come instead from the virtual world". A similar concept is that of "engagement", which, in some cases, "it refers to cognitive operations such as attention, effort or agency when performing a task, while in others it refers more generally to participation in activities" (Richardson et al., 2020, p. 1). Then we have the concept of "presence", a term that appears frequently in studies on virtual reality (Baños et al., 2004) and is often associated with Sherry's (2004) description of it as the feeling of being in a mediated space different from the reality in which the person is. As Cavallo and Fryer (2022, p. 60-61) point out, "presence encapsulates the extent to which any medium induces feelings of

engagement, immersion, and transportation into the virtual world. It occurs when the user is unaware of their experience as being mediated". Therefore, presence could be summarized as "the feeling of being there" (Heeter, 1992). On the other hand, Iturregui-Gallardo (2019, p. 54) supports the view that "the terms 'entertainment' and 'enjoyment' can be used as synonyms" and he continues saying that "for entertainment to happen, the spectator should feel immersed in the content which provides emotions that will ultimately bring them to a pleasurable state".

As these definitions might hint, there is no clear consensus as to what immersion is. Baños et al. (2004) state that, although there have been some attempts to differentiate between presence and immersion, immersion would be the objective description of the technology, whereas presence would be the user's subjective experience, which means that, if something is immersive, someone would feel presence. Iturregui-Gallardo (2019) associates two components with immersion: flow and presence. The concept of "flow", which was proposed by Csikszentmihalyi (1988) and later applied to media by Sherry (2004), could be defined as "the pleasure experienced from the immersion in different activities and tasks" (Iturregui-Gallardo, 2019, p. 54). As for presence, Lombard and Ditton (1997) define it as the illusion of a mediated experience not being mediated. In this vein, we support Iturregui-Gallardo's (2019, p. 55) statement: "in some way, the concept of presence when applied to audiovisual media refers to the immersion produced by non-real and mediated input and the experience of transportation inside the mediated environment without being physically moved". Therefore, we suggest that when an audiovisual product produces an immersive experience, it will trigger high levels of presence in users.

As we have seen, part of AD research moves towards reception and experimental studies, some of which have explored immersion or presence, and some even using heart rate. Following the studies from Baños et al. (2004), Dillon (2006), Riva (2011) and Fryer (2013), the higher the level of presence is, the bigger the emotions that will be felt. In other words, presence is related to emotional activation or arousal. On the other hand, Richardson et al. (2020, p. 2) state that "changes in heart rate have been linked to increased information processing demands and/or greater mental effort" and, more precisely, "increased heart rate is an indicator of increased effort and serves as an indirect measure of cognitive and emotional engagement". The reason to connect AD with immersion is that an immersive product induces the feelings of flow and presence, among others (Fresno, 2017), so if AD scripts attract and hold users' attention, then they can improve their filmic experience. In other words, knowing the key to increasing presence can be a useful way to increase AD effectiveness (Di Giovanni, 2020). In this way, the aim of AD would not only be breaking with comprehension barriers, but immersive AD could also allow users to have higher enjoyment. In this sense, Ramos Caro (2013) explores the emotional impact in an experimental heart rate study with sighted and visually impaired participants. Her results suggest that there should not be any empirical reason not to use more emotional AD. Fryer and Freeman (2014) carry out one of the first studies on emotions and presence with partially sighted participants by studying different variables: fragments with and without AD, and with human or synthetic voices. Their results suggest that verbal information does not reduce presence, a feeling that would be increased by the use of human voices. Wilken and Kruger (2016) investigate the effect of having (or not having) *mise-en-shot* elements in AD on users' immersion. Following their conclusions, the absence of *mise-en-shot* elements in AD has no influence on the feeling of transportation, but it does on the way the audience

identifies with characters, so these elements might be relevant to immersion. Fresno (2017) studies engagement in AD and suggests that this feeling is triggered by comprehension and immersion. Walczak and Fryer (2017) explore how AD styles and locution can influence the levels of presence.

Regarding the device used for our experiment, this pilot study used the UPTIVO Belt-D, a heart rate sensor that was attached to the participants' chest. While the versatility of the device would allow the experiment to take place anywhere, it was restricted to the Asociación Retina facilities, which was familiar to all participants and a location where they felt at ease. While this circumstance might affect the ecological validity of the experiment, since participants do not usually consume audiovisual products in this situation, the main advantage is the possibility of controlling most variables.

3.3 Materials

This experiment is the continuation of a previous eye-tracking study, where the material analysed was narrowed down to two short fragments of video in order to be able to manage a feasible amount of eye-tracking data (for an in-depth explanation, see Romero-Muñoz, in press). Drawing on those two excerpts, in this experiment, a one-minute fragment from *Guillermo del Toro's Cabinet of Curiosities* (episode 1, season 1) was used as an experiment simulation so that participants could understand the dynamic of the experiment without having prior knowledge related to the actual video that was going to be used for the experiment, a five-minute fragment from *Money Heist* (episode 2, season 3).

Table 1: Fragment from the official and creative ADs

Time in	Time out	Original AD	Creative AD
00.00.00.19	00.00.26.11	Una serie Netflix. Una producción Vancouver Media. En el vestíbulo del Banco de España, Lisboa, Estocolmo y Tokio empuñan sus fusiles. Caminan serenas entre los rehenes. Van vestidas con mono rojo y chaleco antibalas. Río exhala y relaja los hombros. Las mujeres comienzan a desvestirse. Denver aprieta los labios y sonríe a Estocolmo.	Una letra N roja de Netflix se despliega como un abanico en un espectro de colores. Título: "Una serie Netflix". Texto en pantalla: "Una producción Vancouver Media". En un vestíbulo, Lisboa, Estocolmo y Tokio empuñan sus fusiles de asalto. Caminan entre los rehenes. Van vestidas con mono rojo y chaleco antibalas para protegerse. Río exhala tranquilo y baja los hombros. Las mujeres se desvisten con cuidado. Denver aprieta los labios, tenso, y le regala una sonrisa a Estocolmo.
00.00.27.23	00.00.31.12	Río le regala un guiño a Tokio mientras toma el fusil de ella.	Río le guiña un ojo a Tokio mientras toma decidido el fusil de ella.
00.00.32.11	00.00.57.07	Palermo grita a los rehenes "¡Caretas fuera!" y ellos se quitan las máscaras. Río deja las armas de las tres mujeres sobre una pared. A continuación, camina entre las dos hileras de rehenes para que estos vayan echando sus caretas en una bolsa de deporte que sujeta con las dos manos. Las tres mujeres se descalzan y se arremangan las perneras del pantalón.	Palermo grita a los rehenes y ellos se quitan obedientes las máscaras. Río deja las armas de las tres mujeres lentamente sobre una pared. A continuación, camina decidido entre las dos hileras de rehenes asustados para que estos vayan echando sus caretas en una bolsa. Las tres mujeres se descalzan con lentitud y se arremangan cuidadosamente las perneras del pantalón.

Source: Author (2025)

The reason for using just one video for the experiment is the fact that two variables were added to *Money Heist*'s fragment: creativity and gender. In a previous descriptive study carried out by Romero-Muñoz (2023), AD scripts proved to be a mixture of objective and subjective multimodal components, so the original AD script's contents were manipulated in creative terms: objective elements were turned subjective and subjective elements were turned objective (Table 1), drawing on the professional and academic proposals presented in section 2. According to Romero-Muñoz's (2023) study, 80% of the AD contents were objective, which means that the creative AD in this experimental study was 80% subjective.

As for gender, the research aimed to determine whether gender can be relevant to immersion, so the professional studio that recorded the official Netflix AD from *Money Heist* in Spain (Mira lo Que Digo) recorded four different AD versions in Spanish (Table 2): the original AD recorded by a man (Test 1), the original AD recorded by a woman (Test 2), a creative AD recorded by a man (Test 3), and a creative AD recorded by a woman (Test 4).

Table 2: Heart rate experiment Ads

Tests	Video fragment	AD	Voice
Test 1	*Money Heist*	original	man
Test 2	*Money Heist*	original	woman
Test 3	*Money Heist*	creative	man
Test 4	*Money Heist*	creative	woman

Source: Author (2025)

Thus, we consider that our control group is made up of the participants who listened to the original AD, both with the male and female voices, since we believe that the experience of these participants is the closest to what users might have felt when listening to the AD from *Money Heist* in real life. On the other hand, the participants who listened to the creative AD (either with the woman's or man's voice) make up our experimental group, who experienced a far more subjective AD than the real one from Netflix (Table 3).

Table 3: Control and experimental groups

Test	Creativity	Gender	Group
1	original	man	control
2	original	woman	control
3	creative	man	experimental
4	creative	woman	experimental

Source: Author (2025)

3.4 Participants

Regarding participants, it must be stated that we present a pilot study here to verify the validity of the experiment, "this will allow the researcher to test selected methods of analysis and will give a feeling for how much data might need to be collected to establish some level of credibility" (Saldanha & O'Brien, 2013, p. 22). Therefore, we limit this first approach to two participants with certain common characteristics: partial sight. There are many classifications regarding people with some major visual loss: blindness, low vision, vision impairment, partial sight, etc. Moreover, this vision loss can have different degrees (total or partial) and have multiple causes: diabetes, glaucoma,

macular degeneration, cataracts, etc. (see López Rubio, 2024, for an overview). According to ONCE, the Spanish association for the blind and partially sighted, in 2022, 81.7% of its members were partially sighted, whereas 13.7% were blind. Taking into account this numerical predominance of partial sight over total blindness, we contacted a Spanish retinal dystrophy association called Asociación Retina, whose members were partially sighted. Once we had partially sighted participants, they were assigned a code (e.g., P001) to anonymise their results, a fact about which participants have to be warned before proceeding with the experiment. This also follows Saldanha and O'Brien's (2013, p. 162) recommendation: "the researcher ought to apply a unique identifier to each participant in order to identify them, but the connection between the identifier and the participant's real name should be available only to the researcher".

3.5 Procedure

The combination of questionnaires, heart rate data, and interviews allowed the mixing of qualitative and quantitative data so that results could be triangulated, that is, "cross-checking the results one set of data provides with results from another set of data" (Saldanha & O'Brien, 2013, p. 23). Bearing in mind this combination of data collection methods, we applied the following procedure in the pilot study. The researcher guided each partially sighted participant to a private room within the association's facilities, where the experiment took place. Once there, the researcher read the informed consent and the informational document out loud, which was signed by asking the participant if they agreed, a process which was recorded. In the same vein, the researcher read the demographic questionnaire out loud, allowing the participant time to select the appropriate categories. After that, an experiment simulation took place with the aim that the participant could get used to the experiment's dynamic. First, the heart rate sensor was placed around the participant's chest, but it was not activated yet. Then, the participant watched a short video fragment from the Netflix series *Guillermo del Toro's Cabinet of Curiosities*, and after that a short version of the presence and the comprehension questionnaires were presented so that the participant could anticipate what sort of questions they were expected to answer afterwards. Having completed the simulation, the experiment as such started. The heart rate sensor was activated and synchronized via Bluetooth with the UPTIVO app, then the researcher left three minutes for relaxation so that the participant's heartbeat could go back to their basal heart rate, after which the sensor started to compile data. The participant was randomly exposed to one of the four *Money Heist*'s ADs (Test 1, 2, 3, or 4), after which the heart rate sensor stopped compiling data. The researcher then allowed the participant three minutes of relaxation, after which the measurement was written down. After that, the participant was given the full presence and comprehension tests, followed by the semi-structured interview. Throughout the whole process, the participant never knew about the exact aim of the experiment in order to minimize the Hawthorne effect, which happens "when people alter (usually improve) their normal behavior because they are aware that they are being studied" (Saldanha & O'Brien, 2013, p. 31). The summary of the experiment procedure is described in Table 4.

Table 4: Heart rate experiment procedure

1. Informed consent
2. Informational document
3. Demographic questionnaire
4. Experiment simulation:
- Heart rate sensor placement
- Short clip from *Guillermo del Toro's Cabinet of Curiosities*
- Short presence questionnaire
- Short comprehension questionnaire
5. Heart rate experiment:
- Heart rate activation
- Three minutes of relaxation (basal heart rate)
- Heart rate data compiling
- Full clip from *Money Heist* with:
▪ Official AD voiced by a man (Test 1)
▪ Official AD voiced by a woman (Test 2)
▪ Creative AD voiced by a man (Test 3)
▪ Creative AD voiced by a woman (Test 4)
- Heart rate sensor stops recording
- Three minutes of relaxation (heart rate after relaxation)
- Heart rate sensor deactivation
- Full presence questionnaire
- Full comprehension questionnaire
6. Semi-structured interview

Source: Author (2025)

4. Analysis and preliminary results

As we have stated, our aim was to design a heart rate experiment to study the combination of objective and subjective multimodal components in AD that are best received by end-users in terms of immersion. The UPTIVO Belt-D app retrieved several data values in beats per minute (bpm): average heart rate, maximum heart rate, and minimum heart rate. Beyond this, we also included two other measurements: basal heart rate and heart rate after relaxation. Moreover, this sensor retrieves bpm every five seconds so that participants' data can be monitored throughout the experiment, not just afterwards. Regarding the principles that direct what multimodal contents to consider in creative terms, this research stems from previous studies where Romero-Muñoz (2023, 2025) retrieved the multimodal components that comprise AD scripts from a descriptive viewpoint, a combination labelled as "multimodal configuration". Romero-Muñoz (2023, 2025) used Chaume's (2004) multimodal proposal for meaning codes (constituted by semiotic signs), some of which are transmitted by means of the visual channel (iconographic, photographic, mobility, shot, graphic, and editing codes) and others by the acoustic channel (linguistic, paralinguistic, musical, special effects, and sound position codes). A corpus of AD scripts was operationalized by using qualitative content analysis, it was classified in creative terms resorting to certain objectivity and subjectivity parameters based on professional and research proposals mentioned in section 2, and it was determined that the multimodal configuration of AD tends to revolve around four sources of information: references to movement, to iconography, to spatial-temporal changes, and to graphic elements on screen. Moreover, according to these results, objectivity and subjectivity seem to coexist in AD scripts, particularly in reference to movement, iconography, and spatial-temporal changes.

As mentioned previously, partially sighted participants were randomly assigned one of the four tests, two of which included the original AD script (one voiced by a man and one voiced by a woman) while the remaining two included the creative AD script (again with male and female voices). Since this is a pilot study that seeks to assess the viability of the research design, we have limited results to the data from two participants with opposite variables: P007 and P010. Participant P007 was a man and he was assigned Test 1 (the original *Money Heist* AD voiced by a man), whereas participant P010 was a woman and she was assigned Test 4 (the creative *Money Heist* AD voiced by a woman). We must remember that Tests 1 and 2 had the original AD and their contents were 80% objective, whereas Tests 3 and 4 had the creative AD with 80% of their contents being subjective. Participant P007 had a basal heart rate of 71 bpm, an average heart rate of 67 bpm, a maximum heart rate of 72 bpm, a minimum heart rate of 64 bpm, and a heart rate after relaxation of 67 bpm. Participant P010 had a basal heart rate of 87 bpm, an average heart rate of 91 bpm, a maximum heart rate of 99 bpm, a minimum heart rate of 82 bpm, and a heart rate after relaxation of 85 bpm (Table 5).

Table 5: Participants' overall data

Participant	Gender	Test	Basal bpm	Average bpm	Maximum bpm	Minimum bpm	Bpm after relaxation
P007	man	1	71	67	72	64	67
P010	woman	4	87	91	99	82	85

Source: Author (2025)

Beyond this data on the overall performance from both participants, we needed to know participants' physiological reactions to different multimodal components in creative terms (Table 6). We chose two fragments (sections of five seconds, which is the minimum span of time that the UPTIVO Belt-D detects) for every component of the multimodal configuration: two fragments with references to movement, two fragments containing iconography, two fragments with spatial-temporal changes, and two fragments with graphic elements. Moreover, these two fragments included an objective and a subjective multimodal reference. We must bear in mind that Tests 1 and 4 were radically different in creative terms: all objective references in Test 1 are subjective in Test 4 (and vice versa). In this vein, out of the possible data that could be retrieved from P001, we selected a first objective mobility fragment (that is, a section where the AD has an objective reference to movement) and a second subjective mobility fragment (where the AD has a subjective reference to movement), and so on (this could not be maintained in the graphic code, so both fragments are objective for Tests 1 and 2, while subjective for Tests 3 and 4).

In this way, we were able to analyze some examples, which will be limited here to the intervals from the mobility code due to space restrictions. The first fragment from the mobility code is found in the interval from second 40 to second 45 and it refers to the objective sign *camina* ("he walks") in the original AD, a piece of information with no interpretation whatsoever. In this same interval from the creative AD, the subjective version of that sign is *camina decidido* ("he walks confidently"), where there is an interpretation of the way he walks.

Table 6: Analysed intervals with the heart rate sensor

Code	Fragment	Interval	AD	Obj./Subj.	Test
mobility	1	00:40 – 00:45	original	objective	1-2
			creative	subjective	3-4
	2	00:25 – 00:30	original	subjective	1-2
			creative	objective	3-4
iconographic	1	00:05 – 00:10	original	objective	1-2
			creative	subjective	3-4
	2	01:20 – 01:25	original	subjective	1-2
			creative	objective	3-4
editing	1	03:45 – 03:50	original	objective	1-2
			creative	subjective	3-4
	2	00:00 – 00:05	original	subjective	1-2
			creative	objective	3-4
graphic	1	02:50 – 02:55	original	objective	1-2
			creative	subjective	3-4
	2	03:40 – 03:45	original	objective	1-2
			creative	subjective	3-4

Source: Author (2025)

Next, the second fragment from the mobility code includes a subjective sign in the interval from second 25 to 30 in the original AD, *Río le regala un guiño a Tokyo* ("Río gifts Tokyo with a wink"), where "gifts Tokyo with a wink" is considered metaphorical and highly interpretative. In this interval from the creative AD, that sign was described objectively as *Río le guiña un ojo a Tokyo* ("Río winks at Tokyo"), a denotational description of the facial movements. Having all this data in mind, we can now clarify what our variables are (Table 7).

Table 7: Variables

Variable	Content of the variable	Type of variable
creativity	creative AD original AD	independent
gender	male voice female voice	independent
heart rate metrics	average heart rate maximum heart rate minimum heart rate basal heart rate heart rate after relaxation fragments 1 and 2 (mobility) fragments 1 and 2 (iconographic) fragments 1 and 2 (graphic) fragments 1 and 2 (editing)	dependent
questionnaires	comprehension presence	dependent
interviews	semi-structured interview	dependent

Source: Author (2025)

With regard to the results obtained, since this is just a pilot study, further quantitative data will be left aside, since proper quantitative results require an appropriate number of participants where statistical analyses could be performed. As far as questionnaires are concerned, P007 scored 3.7/5 points on the presence test and 7/10 on the comprehension test, whereas P010 scored 4.3/5 points on the presence test and 8/10 points on the comprehension test. Regarding the interviews,

P010 more clearly summarized the basic contents of the scene, P007 struggled more with the identification of the multimodal contents, both felt the fragment was immersive, they would both approve of either the male or female AD voice, they both thought the AD was objective enough and would not change anything in particular, and both P007 and P010 considered that AD should be like the one they had just heard (not more objective or subjective).

5. Discussion and conclusions

According to the preliminary results obtained in this pilot heart rate study involving two partially sighted participants, there seems to be some overall differences, since P010 shows higher bpm levels than P007. At this stage of the study, it is impossible to ascertain whether these differences are due to the participants' gender, to the gender of the voice, or to creativity (the differences between objectivity and subjectivity in Tests 1 and 4). As Table 7 shows, further quantitative data could be extracted: bpm during fragments 1 and 2 from every code, and within those fragments, bpm in every test (1-4). However, the aim of this proposal is to create and assess the experimental pilot study using a heart rate sensor to measure partially sighted participants' response to different combinations of subjective and objective multimodal components within the AD script. Therefore, the quantitative data presented here should be seen as an evaluation of how heart rate data could be extracted in a future study with an appropriate number of participants and a statistical interpretation of the results.

That being said, beyond quantitative results, some interesting tendencies can be extracted from the questionnaires and interviews. Even if participants' overall data (Table 5) indeed point to Test 4 (more subjective and voiced by a woman) as the one with higher bpm, quantitative data from two participants is not a reliable source of information. However, it must be noted that P010 also had higher levels of presence and comprehension, which aligns with the studies carried out by Ramos Caro (2013), Wilken and Kruger (2016) as well as Walczak and Fryer (2017), where they conclude that there is no reason not to use subjectivity in AD, since subjectivity is well received among participants. As a matter of fact, subjectivity might increase the levels of presence following our limited quantitative data, which should be further analysed with more participants.

Nevertheless, if we triangulate these results with interviews, the qualitative data points to an issue: participants do not have a clear understanding of creativity in AD. Although P007 and P010 had opposite ADs in terms of creativity, they both felt that their fragment was immersive, they both thought that their AD was objective (which was only true in the case of P007), neither of them would change anything in particular, and they affirmed that AD should be like the one they had just heard (not more objective nor more subjective). This inconsistency has been documented before, proof of which is ADLAD (2013), where 71% of the participants felt inclined to include evaluative adjectives in AD, but 54% of them claimed to be against subjective or interpretative AD. In this vein, Mazur and Chmiel (2016, p. 102) state that one of the disadvantages of reception studies is that "respondents themselves may sometimes have insufficient exposure to AD, and they may not know the essence of AD". Therefore, our opinion is that the key is to opt for experimental reception studies where data from questionnaires, physiological measurements, and interviews can be triangulated.

In conclusion, experimental reception studies seem to be making inroads in AVT and MA research. As far as AD is concerned, guidelines covering AD contents have been tackled from prescriptive and descriptive points of view, but rarely using experimental reception studies that allow an empirical source of information about the main AD audience preferences. Consequently, this pilot study proposes a way of analyzing how objective and subjective multimodal contents are received by visually impaired users, even if our results are limited because of our sample. Preliminary data points towards differences between participants, likely due to creativity, even if users do not have a clear idea of what creativity in AD means. Irrespective of these divergences, further experimental reception studies with a considerable number of participants and statistical data need to be implemented among the main audience of AD so as to address their real needs. Future research following this path could provide empirical evidence about the immersive combination of objective and subjective multimodal components in AD to be considered by quality standards.

References

ABNT. (2018). *NBR 16452. Acessibilidade na comunicação: audiodescrição*. ABNT (Associação Brasileira de Normas Técnicas).

ADLAB. (2013). *Report on testing*. Report no. 3, ADLAB (Audio Description: Lifelong Access to the Blind) Project.

AENOR. (2005). *Norma UNE 153020. Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías* [The Standard UNE 153020. Audio description for visually impaired people. Guidelines for audio description procedures and for the preparation of audio guides]. AENOR.

Audio Description Coalition. (2009). *Standards for Audio Description and Code of Professional Conduct for Describers Based on the Training and Experience of Audio Describers and Trainers from across the United States*. Audio Description Coalition.

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and emotion: their impact on the sense of presence. *CyberPsychology & Behavior, 7*(6), 734–741. https://doi.org/10.1089/cpb.2004.7.734

Bardini, F. (2020). *La transposició del llenguatge cinematogràfic en l'audiodescripció i l'experiència fílmica de les persones amb discapacitat visual* [Cinematic language transposition in audio description and visually impaired people's filmic experience]. [Doctoral Dissertation]. Universitat de Vic. https://www.tesisenred.net/handle/10803/669901#page=1

Cabeza-Cáceres, C. (2013). *Audiodescripció i recepció: efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica.* [Doctoral Dissertation]. Universitat Autònoma de Barcelona. http://hdl.handle.net/10803/113556

California Audio Describers Alliance. (2009). *Audio Description Standards*. California Audio Describers Alliance.

Cavallo, A., & Fryer, L. (2022). *Integrated Access in Live Performance*. Routledge.

Chaume, F. (2004). *Cine y Traducción*. Cátedra.

Csikszentmihalyi, M. (1988). The Flow Experience and its Significance for Human Psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal experience: Psychological Studies of Flow in Consciousness* (pp. 15–35). Cambridge University Press.

Delgado, H., Matamala, A., & Serrano, J. (2015). Speaker Diarization and Speech Recognition in the Semi-automatization of Audio Description: An Exploratory Study on Future Possibilities? *Cadernos de Tradução*, *35*(2), 308–324. https://doi.org/10.5007/2175-7968.2015v35n2p308

Diaz Cintas, J., & Szarkowska, A. (2020). Experimental Research in Audiovisual Translation – Cognition, Reception, Production. *The Journal of Specialised Translation*, (33), 3–16. https://doi.org/10.26034/cm.jostrans.2020.542

Di Giovanni, E. (2014). Visual and Narrative Priorities of the Blind and Non-blind: Eye-tracking and Audio Description. *Perspectives: Studies in Translatology, 22*(1), 136–153. https://doi.org/10.1080/0907676X.2013.769610

Di Giovanni, E. (2020). Reception Studies and Audiovisual Translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 397–413). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-42105-2_20

Dillon, C. (2006). *Emotional Responses to Immersive Media.* University of London.

Fels, D. I., Udo, P. J., Diamond, J. E., & Diamond, J. (2006). A Comparison of Alternative Narrative Approaches to Video Description for Animated Comedy. *Journal of Visual Impairment & Blindness*, *100*(5), 295–305.

Frazier, G. (1975). *The Autobiography of Miss Jane Pittman: An All-Audio Adaptation of the Teleplay for the Blind and Visually Handicapped*. [Master's Thesis]. San Francisco State University.

Fresno, N. (2017). Approaching Engagement in Audio Description. *Rivista Internazionale di Tecnica della Traduzione,* (19), 13–32. https://doi.org/10.13137/2421-6763/17349

Fryer, L. (2013). *Putting It into Words: The Impact of Visual Impairment on Perception, Experience and Presence.* [Doctoral Dissertation]. University of London. https://research.gold.ac.uk/id/eprint/10152

Fryer, L., & Freeman, J. (2012). Cinematic Language and the Description of Film: Keeping AD Users in the Frame. *Perspectives: Studies in Translatology, 21*(3), 412–426. https://doi.org/10.1080/0907676X.2012.693108.

Fryer, L., & Freeman, J. (2014). Can You Feel What I'm Saying? The Impact of Verbal Information on Emotion Elicitation and Presence in People with a Visual Impairment. In A. Felnhofer & O. D. Kothgassner (Eds.), *Challenging Presence: Proceedings of the 15th International Conference on Presence* (pp. 99–107). WUV Universität Wien.

González, C., & Jankowska, A. (2024). Measuring Psychological Immersion through Cardiovascular Response Measures in Subtitled Films. *Translation, Cognition & Behavior, 7*(1), 90–115. https://doi.org/10.1075/tcb.00090.gon

Greco, G. M., & Jankowska, A. (2020). Media Accessibility Within and Beyond Audiovisual Translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 57–81). Palgrave Macmillan.

Hasegawa-Johnson, M. A., Black, A. W., Ondel, L., Scharenborg, O., & Ciannella, F. (2017). Image 2 Speech: Automatically Generating Audio Descriptions of Images. In K. Smaïli, M. Diouri & K. Benzakour (Eds.), *Proceedings of the International Conference on Natural Language, Signal and Speech Processing* (pp. 1–5). ICNLSP. https://hdl.handle.net/2066/244344

Heeter, C. (1992). Being There: The Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments*, *1*(2), 262–271. https://doi.org/10.1162/pres.1992.1.2.262

Holsanova, J. (2016). A Cognitive Approach to Audio Description. In A. Matamala & P. Orero (Eds.), *Researching Audio Description: New Approaches* (pp. 49–73). Palgrave Macmillan.

Hyks, V., & Blokland, B. (2023). *European Blind Union Handbook for High Quality Audio Description on Screen*. European

Iglesias Fernández, E., Martínez, S. M., & Núñez, A. J.C. (2015). Cross-fertilization between Reception Studies in Audio Description and Interpreting Quality Assessment: The Role of the Describer's Voice. In R. Baños-Piñero & J. Díaz-Cintas (Eds.), *Audiovisual Translation in a Global Context* (pp. 72–95). Palgrave Macmillan. https://doi.org/10.1057/9781137552891_5

ITC [Independent Television Commission]. (2000). *Guidance on Standards for Audio Description*. ITC.

Iturregui-Gallardo, G. (2019). *Audio Subtitling: Voicing Strategies and their Effect on Emotional Activation*. [Doctoral Dissertation]. Universitat Autònoma de Barcelona. https://www.tdx.cat/handle/10803/667158

Jankowska, A. (2015). *Translating Audio Description Scripts: Translation as a New Strategy of Creating Audio Description*. Peter Lang.

Jankowska, A., Pilarczyk, J., Wołoszyn, K., & Kuniecki, M. (2022). Enough is Enough: How Much Intonation is Needed in the Vocal Delivery of Audio Description? *Perspectives, 31*(4), 705–723. https://doi.org/10.1080/0907676X.2022.2026423

Krejtz, K., Rutkowska-Siuda, D., & Krejtz, I. (2024). Gaze-Led Audio Description (GLAD). Concept and Application to Accessibility of Architectural Heritage. In A. Marcus-Quinn, K. Krejtz & C. Armando Duarte (Eds.), *Transforming Media Accessibility in Europe* (pp. 53–72). Springer.

Krejtz, I., Szarkowska, A., Walczak, A., Krejtz, K., & Duchowski, A. T. (2012). Audio Description as an Aural Guide of Children's Visual Attention: Evidence from an Eye-tracking Study. In S. N. Spencer (Ed.), *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 99–106). Association for Computing Machinery.

Kruger, J. L. (2010). Audio Narration: Re-narrativising Film. *Perspectives: Studies in Translatology, 18*(3), 231–249. https://doi.org/10.1080/0907676X.2010.485686

Kruger, J. L. (2012). Making Meaning in AVT: Eye Tracking and Viewer Construction of Narrative. *Perspectives: Studies in Translatology, 20*(1), 67–86. https://doi.org/10.1080/0907676X.2011.632688

Kruger, J., & Doherty, S. (2018). Triangulation of Online and Offline Measures of Processing and Reception in AVT. In E. Di Giovanni & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 91–109). John Benjamins.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments, 10*(3), 282–297. https://doi.org/10.1162/105474601300343612

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication, 3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

López Rubio, M. (2024). *Descripción y recepción de la audiodescripción de comedias en Netflix España para personas con discapacidad visual: una propuesta de mejora*. [Doctoral Dissertation]. Universitat de València. https://hdl.handle.net/10550/98663

Mangiron, C. (2022). Audiovisual Translation and Multimedia Localization. In F. Zanettin & C. Rundle (Eds.), *The Routledge Handbook of Translation and Methodology* (pp. 410–423). Routledge.

Matamala, A., Soler-Vilageliu, O., Iturregui-Gallardo, G., Jankowska, A., Méndez-Ulrich, J.-L., & Serrano Ratero, A. (2020). Electrodermal activity as a measure of emotions in media accessibility research: methodological considerations. *The Journal of Specialised Translation*, (33), 192–151. https://doi.org/10.26034/cm.jostrans.2020.551

Mazur, I. (2020). Audio Description: Concepts, Theories and Research Approaches. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 227–247). Palgrave Macmillan.

Mazur, I. & Chmiel, A. (2016). Should Audio Description Reflect the Way Sighted Viewers Look at Films? Combining Eye-tracking and Reception Study Data. In A. Matamala & P. Orero (Eds.), *Researching Audio Description: New Approaches* (pp. 97–121). Palgrave Macmillan.

Morisset, L., & Gonant, F. (2008). *Charte de l'audiodescription. Principes et orientations* [Audio description charter. Principles and orientations]. Ministère des Affaires Sociales. https://www.alain-bensoussan.com/wp-content/uploads/255205.pdf

Naves, S. B., Mauch, C., Alves, S. F., & Araújo, V. L. S. (Orgs.). (2016). *Guia para Produções Audiovisuais Acessíveis*. Ministério da Cultura, Secretaria do Audiovisual.

Netflix (2024). *Audio Description Style Guide (v2.5)*. https://partnerhelp.netflixstudios.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v2-5

Ofcom. (2024). *Ofcom's Guidelines on the Provision of Television Access Services*.

Orero, P., & Vilaró, A. (2012). Eye Tracking Analysis of Minor Details in Films for Audio Description. *MonTi. Monographs in Translation and Interpreting*, (4), 295–319. http://dx.doi.org/10.6035/MonTI.2012.4.13

Orrego-Carmona, D. (2019). Audiovisual Translation and Audience Reception. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 367–382). Routledge.

Pedersen, J. (2020). Audiovisual Translation Norms and Guidelines. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 417–436). Palgrave Macmillan.

Perego, E., & Pacinotti, R. (2020). Audiovisual Translation through the Ages. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 33–56). Palgrave Macmillan.

Ramos Caro, M. (2013). *El impacto emocional de la audiodescripción*. [Doctoral Dissertation]. Universidad de Murcia. http://hdl.handle.net/10201/36475

Ramos Caro, M. (2015). The Emotional Experience of Films: Does Audio Description Make a Difference? *The Translator, 21*(1), 68–94. https://doi.org/10.1080/13556509.2014.994853

Ramos Caro, M. (2016). Testing Audio Narration: The Emotional Impact of Language in Audio Description. *Perspectives, 24*(4), 606–634. https://doi.org/10.1080/0907676X.2015.1120760

Remael, A., Reviers, N., & Vercauteren, G. (2014). *Pictures Painted in Words. ADLAB Audio Description Guidelines*. ADLAB Project.

Richardson, D. C., Griffin, N. K., Zaki, L., Stephenson, A., Yan, J., Curry, T., Noble, R., Hogan, J., Skipper, J. I., & Devlin, J. T. (2020). Engagement in Video and Audio Narratives: Contrasting Self-report and Physiological Measures. *Sci Rep*, (10), 2020. https://doi.org/10.1038/s41598-020-68253-2

Riva, G. (2011). Presence, Actions and Emotions: A Theoretical Framework. *Annual Review of CyberTherapy and Telemedicine, 9,* 2–5.

Rizzo, A., & Spinzi, C. G. (2023). Authorial (Audio) Description: Creativity in the Transfer of CSRs in *Squid Game. Translation and Translanguaging in Multilingual Contexts*, 9(3), 419–449. https://doi.org/10.1075/ttmc.00121.riz

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*, 1–15. https://doi.org/10.3389/FPSYG.2021.661452

Romero-Fresco, P. & Chaume, F. (2022). Creativity in Audiovisual Translation and Media Accessibility. *The Journal of Specialised Translation,* (38), 75–101. https://doi.org/10.26034/cm.jostrans.2022.084

Romero-Muñoz, A. (2023). Multimodal Analysis as a Way to Operationalise Objectivity in Audio Description: A Corpus-based Study of Spanish Series on Netflix. *Journal of Audiovisual Translation, 6*(2), 8–32. https://doi.org/10.47476/jat.v6i2.2023.251

Romero-Muñoz, A. (2025). The Audio Description Script through the Lens of Multimodality: A Qualitative and Quantitative Analysis of the Meaning Codes in *Elite. Parallèles*, 1–15. https://doi.org/10.17462/para.2025.02.03

Romero-Muñoz, A. (in press). Towards an experimental multimodal configuration. A methodological proposal for eye-tracking the audio description script. In F. Chaume & L. Kostopoulou (Eds.), *Translation as Experimentation. Audiovisual and Interactive Media.* (forthcoming).

Saldanha, G., & O'Brien, S. (2013). *Research Methodologies in Translation Studies*. Routledge.

Sherry, J. J. L. (2004). Flow and Media Enjoyment. *Communication Theory, 14*(4), 328–347. https://doi.org/10.1111/j.1468-2885.2004.tb00318.x

Snyder, J. (2014). *The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description*. American Council of the Blind.

Soler Gallego, S., & Luque Colmenero, M. O. (2023). Increased Subjectivity in Audio Description of Visual Art: A Focus Group Reception Study of Content Minimalism and Interpretive Voicing. *Journal of Audiovisual Translation*, *6*(2), 55–76. https://doi.org/10.47476/jat.v6i2.2023.248

Szarkowska, A. (2011). Text-to-speech audio description. Towards wider availability of AD. *The Journal of Specialised Translation*, (15), 142–162. https://doi.org/10.26034/cm.jostrans.2011.509

Szarkowska, A. (2013). Auteur Description: From the Director's Creative Vision to Audio Description. *Journal of Visual Impairment & Blindness*, *107*(5), 383–387. https://doi.org/10.1177/0145482X1310700507

Szarkowska, A., & Jankowska, A. (2012). Text-to-speech Audio Description of Voiced-over Films: A Case Study of Audio Described *Volver* in Polish. In E. Perego (Ed.), *Emerging Topics in Translation: Audio Description* (pp. 81–94). Edizioni Università Di Trieste.

Szarkowska, A., & Mączyńska, M. (2011). *Text-to-speech Audio Description with Audio Subtitling to a Nonfiction Film: A Case Study Based on La Soufriere by Werner Herzog*. University of Warsaw.

Szarkowska, A., & Wasylczyk, P. (2014). Audiodeskrypcja autorska [Auteur audio description]. *Przekładaniec*, *28*, 48–62. https://doi.org/10.4467/16891864PC.14.004.1711

Technical specification ISO/IEC TS 20071-21. (2015). *Information technology, user interface component accessibility. Part 21*. ISO.

Udo, J. P., & Fels, D. I. (2009). From the Describer's Mouth: Reflections on Creating Unconventional Audio Description for Live Theatre. In A. Şerban, A. Matamala & J.-M. Lavaur (Eds.), *Audiovisual Translation in Close-up* (pp. 257–278). Peter Lang.

Walczak, A. (2017). *Immersion in Audio Description: The Impact of Style and Vocal Delivery on Users' Experience*. [Doctoral Dissertation]. Universitat Autònoma de Barcelona]. https://www.tdx.cat/handle/10803/402401#page=1

Walczak, A., & Fryer, L. (2017). Creative Description: The Impact of Audio Description Style on Presence in Visually Impaired Audiences. *British Journal of Visual Impairment, 35*(1), 6–17. https://doi.org/10.1177/0264619616661603

Wilken, N., & Kruger, J. L. (2016). Putting the Audience in the Picture: *Mise-en-shot* and Psychological Immersion in Audio Described Film. *Across Languages and Cultures, 17*(2), 251–270. https://doi.org/10.1556/084.2016.17.2.6

Notes

Authorship contribution

Conceptualization: A. Romero-Muñoz
Data collection: A. Romero-Muñoz
Data analysis: A. Romero-Muñoz
Results and discussion: A. Romero-Muñoz
Review and editing: A. Romero-Muñoz

Research dataset

The research data is part of Alejandro Romero-Muñoz' doctoral dissertation titled: *Hacia una audiodescripción inmersiva. Estudio descriptivo, comparativo, multimodal y de recepción experimental sobre la objetividad y subjetividad en audiodescripción.*

Image copyright

Not applicable.

Approval by ethics committee

This experiment was approved by the Human Research Ethics Committee of the University Jaume I (file number CEISH/76/2023).

Conflicts of interest

Not applicable.

Data availability statement

The data from this research, which are not included in this work, may be made available by the author upon request.

License

Publisher

Guest editors

Willian Moura – Iván Villanueva-Jordán

Section editor

Andréia Guerini

Technical editing

Alice S. Rezende – Ingrid Bignardi – João G. P. Silveira – Kamila Oliveira

Article history