

WORDS FOR VISUAL: ELIMINATING BARRIERS FOR COMMUNICATION

María J. Valero Gisbert¹; Domenico Cosenza²

¹Associate Professor, Department of Humanities, Social Sciences and Cultural Industries, Lingua e Letteratura Straniere Unit, University of Parma, Parma, Italy

²Professor of Psychopathology of Eating Behaviour/Brain System and Behaviour Sciences, Department of Psychology, University of Pavia, Italy

ARTICLE INFO	ABSTRACT
<p>Article History:</p> <p>Received 15.08.2025 Accepted 15.10.2025 Published 15.11.2025</p> <p>Keywords:</p> <p>audio description, visual literacy, accessibility</p>	<p><i>In this research, we focus on audio description (AD) as a form of audiovisual translation designed to eliminate barriers in visual communication for people who, for various reasons, lack access to visual content. To achieve this, it is essential to reflect on visual literacy to understand a world increasingly dominated by image-based communication. This becomes especially relevant when considering the description of images in general. Secondly, we analyse a multimodal text, specifically a film text. Using audio description as a starting point, we aim to reflect on the application of this modality in cinema.</i></p> <p><i>The methodology includes an introduction to AD, its objectives, and the outcomes that result in an accessible version of audiovisual content. For this process, it is necessary to have a foundational understanding of translation, image grammar, and film language, as well as a minimum level of visual literacy. This foundation enables the analysis of the text with the objective of translating it, producing an accessible version, or creating an audio-written script.</i></p> <p><i>Building on these considerations, this research proposes a reflection on existing description practices to evaluate their comprehensibility and the mental images they evoke. Specifically, we investigate what is conveyed through words and what images are constructed in the minds of the audience based on these descriptions. A representative sample or selection of proposed descriptions will be analyzed by comparing different versions of the same film. This comparison will address: 1) The cognitive load (focusing on syntax and lexis); 2) The creation of mental images through words. Regarding the second point, we explore: 2a) Whether the mental images differ for a sighted person, someone born blind, or someone who has lost their sight. 2b) Whether the descriptions can produce the same mental images for audiences from different cultural backgrounds.</i></p>
<p>Copyright©2021 by author. This is an open access article distributed under the Creative Commons Attribution License - Non-Commercial 4.0 International (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.</p>	

1. Introduction

The article 19 of the Universal Declaration of Human Rights¹ (UDHR) intitled ‘Living independently and being included in the community’, recognise “the equal right of all persons with disabilities to live in the community, with choices equal to others, and shall take effective and appropriate measures to facilitate full enjoyment by persons with disabilities of this right and their full inclusion and participation in the community”.

Starting from this premise, and with the awareness that the majority of a country's tangible and intangible cultural heritage remains inaccessible to people with visual sensory disabilities, this research aims to highlight once again the need to provide every individual with the possibility and the freedom to choose cultural and social activities of interest, regardless of their physical and/or intellectual condition.

What ultimately matters, as we have emphasized, and as authors such as Cho et al. point out in relation to individuals with visual impairments, underlining the importance of accessibility, is that this community: “[...] must have access to the world’s visual culture if they are to participate fully in their communities and the world at large. Such access will improve the quality of their lives and help them gain skills crucial to their education and employment opportunities”. Cho et al. (2020: 1).

Our aim in this research is to demonstrate how audio description (AD) can provide accessible audiovisual content, ensuring that everyone has equal opportunities to enjoy cultural products.

2. Translation through Audio description

Audio description (AD) consists of the verbal description of images. It is a mode of translation that can be situated within the field of Audiovisual Translation (AVT) when dealing with audiovisual products. In a broader sense, audio description should be understood as a form of *intersemiotic translation* (Jakobson 1959), and accessible whose aim is to render visible any content that is typically perceived through the sense of sight.

The development of this translation modality and the creation of effective audio descriptions require a high degree of specialization, depending on the nature of the object or product being described. Each type of cultural object presents specific constraints. For instance, in a museum setting, AD can be supplemented by tactile materials or generic information in Braille. In contrast, for audiovisual content, AD must be seamlessly integrated into the film, synchronized precisely with the flow of images and sound.

To create an AD, the translator must not only be trained in translation theory and practice but also, as is particularly relevant in this context, in film language (Casetti and De Chio 1990) and the grammar of visual design (Kress and Van Leeuwen 1996, Dondis 2006). This knowledge is essential for conducting a semiotic analysis of the cinematic text and for being able to verbalise, through language, the various codes present in a multimodal text, in this case, a film. Indeed, as Aiello argues, revisiting the foundations of semiology, particularly the work of Barthes (1961, 1964) and the levels of image interpretation (denotative and connotative), that attention must be paid to the underlying codes that structure visual meaning, and asserts that:

¹ <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-19-living-independently-and-being-included-in-the-community.html>

“Codes can be defined as the ‘implicit rules’ (Sturken and Cartwright, 2009: 434) that govern the ways in which those who make and use images ‘read’ their meanings. As part of a shared system of culture, most of us are able to draw from the same codes to interpret and understand images.” (Aiello 2020: 370).

Moreover, as part of the preparatory training, it is crucial to develop the ability to interpret visual language, what Debes (1969) referred to as *Visual Literacy*. For this author, this concept refers to:

“[...] a group of vision-competencies a human being can develop by seeing and at the same time having and integrating other sensory experiences. The development of these competencies is fundamental to normal human learning. When developed, they enable a visually literate person to discriminate and interpret the visible actions, objects, symbols, natural or man-made that he encounters in his environment. Through the creative use of these competencies, he is able to communicate with others. (Debes, 1969, p. 27)”

The semiotic analysis of the work is essential not only for determining which elements should be described and which can be omitted, but first and foremost, for achieving a proper understanding of the text.

2.1 Visual Literacy

With regard to image analysis for the purpose of description, Avgerinou and Ericson (1997: 286) point out that this process is of utmost importance *“understand and judge visual images”*.

They propose four phases related to the decoding process, which consist of “a) description of the graphic elements composing the image. b) analysis of the ways those elements have been arranged. c) Interpretation of the messages being communicated, and d) aesthetic appreciation of the image.”

The underlying idea, from a broader perspective, is that, whether we acknowledge it or not, we live in a world where images play a predominant role. As a result, it becomes increasingly necessary to provide training that considers modes of communication which, beyond verbal and textual codes, also involve visual elements.

In terms of education, these scholars (1997: 288) question how it is possible that such little emphasis is placed on visual communication in academic curricula. They argue that: *“The fact that we do learn and are influenced by what we see on television, [sic] is beyond doubt; and for a high proportion of children in schools, the influence of visual media in their adult life will be infinitely more dominant than print.”*

These words remain relevant even today, nearly 30 years later, as learning is increasingly mediated by screens (tablets, smartphones, and to a lesser extent, television). One might naïvely argue, we believe, that explicit instruction is unnecessary and that users acquire visual competence simply through exposure. However, we must recall that, just as mere exposure to certain scripts or alphabets does not grant access to their content, simply seeing an image does not equate to understanding or knowing, or at least not fully.

Along these lines, Avgerinou and Ericson (1983: 288) point out, without specific instruction, we must ask to what extent we are passive receivers truly capable of discerning quality, and how such exposure shapes the development of an individual who lacks the necessary keys and competencies to interpret certain codes: *“How far can this empirical venture guarantee that they are not partial and passive receivers of visual messages; that they are able to make a critical selection between the necessary and the unnecessary; and they do distinguish superficial, glamorous, and pseudo-sophisticated messages from the real and valuable one?”*.

On the other hand, it is important to consider not only what images convey, but also what the creator of that content intentionally aimed to communicate. In this regard, it is particularly relevant to highlight the reference to the concept of *intentionality* from Ausburn and Ausburn (1978: 291) in the definition of visual literacy as *“a group of skills which enable an individual to understand and its visuals for intentionally communicating with others.”* since, indeed, behind every image, there is an author.

2.2 Audio description: understanding semiotic code-mixing

Considering the image as a system of signs that conveys meaning, Visual Literacy provides the foundation for identifying those signs (such as shapes, colors, and composition). Semiotic analysis, in turn, is necessary to interpret what they communicate. In other words, it examines which elements make up the visual text and how they relate or combine to produce a specific meaning.

The semiotic analysis of visual and acoustic elements is essential for the work of the audio describer or image translator.

Conducting a semiotic analysis of a text involves understanding the interplay between various codes that together construct meaning. In this context, Aiello (2020: 368) specifies that the aim of semiotic analysis is *“to make the hidden structures, underlying cultural codes, and dominant meanings of such texts both visible and intelligible. In doing so, semiotics is also a powerful instrument for a systematic study and critique of ideology in visual communication.”*

The process of creating an audio description (AD) for cinematic products involves several stages, as outlined by Remael et al. (2014) and Valero-Gisbert (2021). Among these, and as an essential preliminary step, is the need for thorough documentation about the film, its creator(s), and contextual elements. This also includes compiling a glossary of terms relevant to the specific cinematic genre. Having a lexicon of commonly used words associated with both the genre and the specific work is considered a best practice that facilitates the translation task.

Audio description ensures equal access to culture (museums, theatre, cinema, etc.) so that people with visual impairments can participate fully in society, free from exclusion or discrimination.

3. Methodology

This study focuses on audio description (AD) as applied to audiovisual products, specifically a sequence from *Nuovo Cinema Paradiso* (Giuseppe Tornatore, 1988). Set in post-war Italy, the film tells the story of a famous film director originally from Cataldo, a Sicilian

town, who returns to his homeland following the death of his mentor, years after leaving it in his youth.

Describing images is not as straightforward as it may seem, especially when they convey content rooted in a specific culture. In this work, we have chosen to analyze an Italian film that depicts European life more than half a century ago, in order to highlight the contrast with more distant cultures. From a translation perspective, our aim is to reflect not only on what these images communicate, but also on what they might suggest in the absence of audio description for culturally distant audiences. This, in turn, underlines the importance of audio description for all viewers.

We have selected² a culturally charged fragment that represents a point of intersection between cinema and the Catholic Church³. Through these images, we are transported into the sociocultural reality of 1950s–60s Italy, reflects particular customs and social conventions.

In this particular sequence, the Church's moral concerns are embodied in the figure of a priest who censors scenes considered immoral at the time, specifically, images of actors kissing. The parish priest of Cataldo reviews the film reels and signals the scenes to be removed by ringing a small bell. The sound of the bell plays a central narrative and symbolic role, becoming a clear marker of censorship.

In this case, the role of the AD is to describe the actions neutrally, without providing additional commentary on the practice of censorship itself, which was common at the time.

We would like to remind that, as Chion explains (1994: 216), “*a kind of symbolic contract that the audio-viewer enters into, consenting to think of sound and image as forming a single entity*”.

The importance of sound, and music in particular, as a transmitter of culture is emphasized by Poulakis y Tzamtzi (2023: 148), who state that: “*Since the beginning of the 21st century, new digital technologies and multimedia practices, combining texts, sounds, and moving images have encouraged several ethnomusicologists and anthropologists of music to create fresh audiovisual representations of musiccultures and disseminate them all over the world.*”

3.1 Blind people and the spatial image: from Diderot to Lacan

On the one hand, there is the need to understand the images; on the other, with the goal of creating mental images, it is important to remember that in an audiovisual product, this will be achieved through the auditory channel. In this context, spatial references provided in the description, along with the specific use of sound (such as using echoes or reverberation to suggest depth), will play a crucial role. In this regard, we bring to mind a reflection by Lacan (1964) on this very aspect.

The methodology we propose is grounded in an idea that was already recognized during the Enlightenment Age by Diderot in his renowned *Lettre sur les aveugles à l'usage de ceux qui voyent* [Letter on the Blind for the use of those who see] from 1749 and has since been systematically studied in later periods: the condition of blindness does not imply a damaged

² The sequence was extrapolated with the software Avidemux 2.8 VC.

³ To that regard, see G. Convents (2001). *I cattolici e il cinema*. In G.P. Brunetta, *Storia del cinema mondiale*, V, *Teoria, strumenti, memorie*, Torino, (pp. 485-517).

relationship with spatial perception or with the internal laws that govern spatial organization. As Lacan stated in his Seminar XI commenting the Diderot's letter, *"...for the geometrical space of vision – even if we include those imaginary parts in the virtual space of the mirror [...] – is perfectly reconstructible, imaginable by a blind man"* (Lacan, 1964, Sem. XI, p. 86 English edition). The Diderot's letter *"shows to what extent the blind man is capable of taking account of reconstructing, imagining, speaking about everything that vision yields to us of space"* (Lacan, 1964, Sem. XI, p. 92).

The *geometrical organization of space in the blind remains intact*, and with the aid of touch (as in museums), a blind person is able to orient themselves and interact with objects in space, even when they cannot see them (Lacan, Sem. XI, p. 93). These premises lay the foundational conditions for a methodology aimed at supporting blind users through audio description, when combined with tactile perception, this approach enables access to cultural products in which visual imagery plays a central role.

4. Audio description versions

We now present three audio-described versions of this scene, two in Italian and one in Spanish. We will analyse similarities and differences, focusing primarily on syntax and lexical choices. These two linguistic dimensions provide insight into: 1) the cognitive load implied by sentence structure and the amount of information presented; 2) the mental imagery constructed through language, imagery that should ideally resemble what a sighted person sees, for a visually impaired listener relying solely on the description. We also ask: 2a) Will this mental image differ for someone who is sighted versus someone who is congenitally blind or has lost their sight later in life? and 2b) can these descriptions generate the same mental image across audiences from different cultural backgrounds?

From a linguistic and translation perspective, careful lexical choices are essential in crafting effective AD. Syntactically, it is important to use short sentences that each convey a complete unit of meaning. This approach reduces cognitive load, ensuring clarity and ease of understanding. Additionally, attention must be paid to the order of elements in each sentence, whether it follows a coherent, cohesive structure that facilitates immediate comprehension (Braun, 2011), and whether some structures may be clearer or more effective than others depending on the audience.

Table 1 – Audio Descriptions of the censure clip

Audio descriptions	AD Scripts
AD Spain ONCE ⁴	The priest enters an empty movie theater and signals to a woman to turn off the lights. He shouts toward the projection booth. The priest sits down, holding the altar boy's handbell. The woman closes the shutters. Alfredo, the projectionist, prepares the film in the bluish dimness of the booth. On the screen, a curly-haired blonde girl embraces Jean

⁴Spanish Audio Description, ONCE 2000. *Nuovo Cinema Paradiso*, 1989. Color, rated suitable for all audiences. Audio description script using the Audesc system written by Rosa Pujol, audio produced at Estudios Alisea, with technical coordination of the system by Javier Navarrete, under the Department of Culture and Sports of ONCE.

	Gabin, they kiss. Totò laughs, and the priest rings the bell. Alfredo places a slip of paper on the reel, marking the segment to be cut. ⁵
AD Trento ⁶	The priest enters the cinema hall. He takes a small bell from his pocket and sits down. Alfredo finishes adjusting the projector. The two actors kiss. At the sound of the bell, Alfredo slips a piece of paper into the film reel without stopping the projection. The ringing of the bell amuses the young boy and irritates Alfredo, who has to insert yet another slip of paper to mark the censored scene. ⁷
AD Torino ⁸	He enters making the sign of the cross in a movie theater and gestures to a cleaning woman to leave. Turning toward the projectionist, whose face appears framed by the window of the projection booth, the priest takes a seat in the audience. Shots of the attentive priest with the bell in hand alternate with images of Alfredo in the booth and Totò. The priest watches the film closely. His fingers tap on the bell. Alfredo smiles as the indignant priest raises his arm and rings the small bell. Totò laughs with his mouth wide open from behind the curtains. Alfredo slips a piece of paper between the winding loops of the reel as a marker. ⁹

Source: Excerpts from the authors of the paper

We can compare the three versions in terms of clarity, level of detail, coherence, and accessibility. However, these criteria alone are not sufficient, since we are dealing with a moving-image text that integrates multiple semiotic codes into a single whole. For instance, dialogue, character sounds, and noises from objects must also be considered when creating an audio description. To support the identification and interpretation of these elements, we have proposed a multimodal transcription framework (Valero-Gisbert 2021). Based on the analyzed AD scripts, we can draw several key considerations, outlined below.

4.1 Similarities and Differences

All three audio descriptions convey a consistent sequence of events: a priest enters an almost empty cinema where a woman is present, either cleaning or turning off the lights. The priest, carrying a small bell symbolising censorship or warning, sits down as Alfredo operates the projector. A kissing scene featuring Jean Gabin appears on screen, prompting the priest to

⁵Original version: El cura entra en una sala de cine vacía, indica a una mujer que apague las luces. Grita hacia el operador. El cura se sienta con la campanilla del monaguillo en la mano. La mujer cierra las persianas. Alfredo, el operador, prepara la película en la penumbra azulada de la cabina. En la pantalla, una chica de rizado cabello rubio se estrecha contra Jean Gabin, se besan, Totò ríe y el cura agita la campanilla. Alfredo coloca un papelillo en la bobina señalando el fragmento qua habrá de cortar.

⁶ **Italian Audio Description.** Franco Cristaldi presents a French-Italian co-production.

⁷ Original version: Il parroco entra nella sala di un cinema. Il prete prende dalla tasca un campanello e si siede. Alfredo finisce di sistemare la macchina da proiezione. I due attori si baciano. Al suono del campanello, Alfredo infila un pezzo di carta nel rotolo di pellicola senza fermare la proiezione. Lo scampanello suscita l'ilarità del ragazzino e il malumore di Alfredo che deve inserire un altro foglietto nel pezzo da censurare.

⁸**Beyond Vision: The Museum to Touch, The Cinema to Listen To. The AD** was made possible thanks to the support of the **National Cinema Museum**, the **Piemonte Region**, the **City of Turin**, and the **CRT Foundation**. Script: **Pieranna Pinna**. Voice: **Simone Schinocca**.

⁹ Original version: Poi entra facendo il segno della croce in una sala cinematografica, fa cenno di uscire ad una donna che pulisce. Rivoltosi all'operatore, il cui volto è apparso incorniciato dalle aperture della cabina di proiezione, il prete si siede in platea. Le immagini del prete attento con campanello in mano si alternano a quelle di Alfredo in cabina e di Totò. Il prete segue attento il film. Le dita tamburellano sul campanello. [sul bacio] Alfredo sorride mentre il prete indignato solleva il braccio e fa suonare il campanellino. Totò ride a bocca aperta dietro le tende. Alfredo inserisce a guisa di segno tra le spire della bobina in avvolgimento un foglietto di carta.

ring the bell. Alfredo marks the scene for removal by inserting a slip of paper into the reel, while Totò responds with amusement.

Although the core content remains the same across descriptions, differences emerge in syntax and vocabulary. The Spanish AD adopts a more objective and neutral tone. The Trento AD is more concise, whereas the Turin AD is the most detailed and the least objective of the three.

From a syntactic perspective, both the Spanish and Trento descriptions use short, self-contained sentences, with the exception of the final one in Trento, which contains a subordinate clause. The Turin version is syntactically more complex, featuring hypotactic structures.

In terms of describing the character's state of mind, the Spanish AD refers only to the child's reaction. The Trento version expands on this by adding emotional cues for the projectionist, using adjectives such as 'hilarity' for the child and 'bad mood' for the operator. The Turin AD, however, adds the priest's emotion 'indignant' which contrasts with Trento, as Alfredo is described as smiling rather than irritated, and Totò is said to laugh.

As for lighting (Maszerowska 2012) and color cues, only the Spanish AD includes a reference to the bluish glow from the projection booth, although it omits the explicit contrast with the dark cinema. This contrast can, however, be inferred from the woman's action of closing the shutters.

Table 2: AD key differences

Aspect	Spanish AD	Trento AD	Turin AD
Clarity/Tone	Objective, neutral	Concise	Most detailed, least objective
Syntax	Short, self-contained sentences	Short sentences, except final one with subordinate clause	More complex, uses hypotactic (subordinate) structures
Focus on emotions	Only mentions Totò's amusement	Adds emotional states: Totò with <i>hilarity</i> , Alfredo in a <i>bad mood</i>	Expands further: Priest is <i>indignant</i> , Alfredo is <i>smiling</i> , Totò is <i>laughing</i>
Level of detail	Minimal, sticks to factual description	Moderate, with some emotional cues	Richest in detail, includes contrasting emotional perspectives

4.2 Cognitive load

As previously explained, the Turin AD presents a complex syntactic structure, mainly composed of subordinate clauses. For instance, the sentence "Rivoltosi all'operatore, il cui volto è apparso incorniciato dalle aperture della cabina di proiezione, il prete si siede in platea" demands greater attention from the audience, thereby resulting in a higher cognitive load.

From a lexical standpoint, the use of common vocabulary such as 'sala de cine' instead of 'cinematográfica', 'campanilla', 'bobina', 'cura', 'párroco' o attori' facilitates the flow of the description. In contrast, the inclusion of less common or specialized terms like 'a guisa de', 'tamburellare', or 'spira' in the phrase "inserisce a guisa di segno tra le spire della bobina in avvolgimento un foglietto di carta" may interrupt the flow, diverting attention to an unfamiliar term.

Word order also affects processing speed. For instance, in the Spanish AD, the description of the girl on screen reads: “chica de cabello rizado rubio”. This is an unusual word order, more typical would be “chica rubia de/con el pelo rizado”, which would improve the fluency and ease of understanding.

In the Trento AD, there is inconsistency in terminology, such as the alternation between ‘prete’ and ‘parroco’ to refer to the same character. To avoid confusing the audience, it's advisable to maintain terminological consistency. Similarly, in the Turin AD, the term ‘campanello’ is later replaced by ‘campanellino’. Though minor, these changes can affect immediate comprehension and the identification of the object.

Another important factor is the amount of information provided. The word count varies significantly between the three versions: ONCE (Spanish): 89 words, Trento: 70 words, Turin: 112 words¹⁰. This directly impacts the speed of narration and requires greater attentional resources from the receiver to process and retain the information, thereby increasing the cognitive load. It is also necessary to allow pauses so that listeners can process the information they hear.

To reduce cognitive load, the audio describers should use clear language, short and complete sentences, consistent terminology, and a natural word order. For example, instead of phrasing it as in the Spanish AD “*On the screen, a curly-haired blonde girl embraces Jean Gabin, they kiss. Totò laughs, and the priest rings the bell. Alfredo places a slip of paper on the reel, marking the segment to be cut.*”, a good practice is to simplify syntactic structures and reduce word count, as follows: “*On screen, a blonde girl kisses Jean Gabin. Totò laughs. The priest rings the bell. Alfredo marks the reel with paper*”.

4.3 Inconsistencies / redundancies / omissions between Descriptions and Images

As we have discussed earlier, the description must consider the interrelation of different codes. Sounds are part of the acoustic code, so they should not be included in the audio description script unless there are overlapping sounds or when the sound does not clearly correspond to the visible action. For example, in the Spanish AD, Totò is described as smiling, but since the smile is silent and no corresponding sound is heard, the action cannot contradict what is heard. The Spanish AD also says the priest “*indicates to a woman to turn off the lights*”, yet the viewer hears the sound of the woman closing the windows, which does not correspond to the action of switching off the lights. In the Turin AD, this instruction is omitted and instead it describes the priest as “*gesturing to a woman who is cleaning to leave*”. The sound of the windows closing is not described or contextualized anywhere, so a blind viewer would not know what causes this sound or be able to match it with any image.

Another important aspect to consider is **redundancies**, meaning information that the viewer can already understand from other cues and therefore should not be repeated in the audio description script (ADs). For example, the ONCE Spanish AD says: “*The priest enters an empty cinema hall, indicates to a woman to turn off the lights. He shouts towards the operator. The priest sits down holding the altar boy's bell. The woman closes the blinds. Alfredo, the operator, prepares the film in the bluish half-light of the booth. On the*

¹⁰ For timing measurement, since all three versions have the same duration, a higher word count necessarily entails a faster narration speed.

screen, a girl with curly blonde hair cuddles Jean Gabin, they kiss, Totò laughs, and the priest shakes the bell [...]”. Clearly, the verb ‘shouts’ should be omitted because it implies a sound: “shouts towards the operator”, the dialogue already conveys this: “Alfredo! Alfredo! You may start”, a better choice would be ‘addresses’ instead of ‘shouts.’ Another element is on “the priest shakes the bell” that is redundant since the sound is already audible.

Regarding information about space and non-verbal language, only the Spanish AD explicitly mentions that the hall is empty, and only the Turin AD notes the priest making the sign of the cross (a Christian ritual gesture) upon entering.

4.4 Mental Images

As can be inferred from the points above, these differences, affecting the staging, the characters’ emotional states, and information about lighting or color, will influence the mental image evoked by the description. Emotions are expressed through actions (tapping fingers, laughing, inserting markers), allowing listeners to infer tone.

It is important to consider that sounds serve as the primary means of accessing content for a person who is blind. Beyond dialogue, here we have the sound of the bell, laughter, and the noise of the projection booth. In the future, this could potentially be enhanced with vibrations depending on the film, or even olfactory elements, to offer a fuller experience for visually impaired audiences.

On the other hand, for someone blind from birth, the image suggested by descriptions like “*voltosi[...] incorniciato dalle aperture della cabina di proiezione* [turned around [...] framed by the openings of the projection booth]” may be unclear, and the other details will result in different mental images depending on the level of detail: physical descriptions of characters, costumes, or other elements of the staging, to name a few¹¹.

4.5 The Same Mental Image Across Different Cultures

The symbolic meaning conveyed by this sequence is likely recognizable in cultures close to the Spanish or Italian, where the Church has played and continues to play a fundamental role in society. The figure of the priest could be interpreted either as oppressive or protective. In the first case, in his role as censor, this might not be as clear in cultures that are more distant or follow different religions, or that role might be taken by another institution. This raises the question of how audio description would be adapted for an audience whose cultural context does not assign the same meaning to the actions shown in the images as the original intended audience, in this case, the Italian audience (Valero-Gisbert, 2024).

A second point concerns the identification of the famous actor Jean Gabin, an icon of 20th-century French cinema, who may not evoke the same emotions or associations in other cultures.

Additionally, contemporary cinema has lost some of the social function depicted in the film, as a meeting place for all social class, which may be less comprehensible today, especially in other cultural contexts.

¹¹ A reception study on mental imagery will be addressed in future research, as it requires cross-cultural participation to validate the hypothesis.

Ultimately, since cultural values and references are not universal, as demonstrated here, the impact of images will never be the same everywhere, and interpretation will always be influenced by our own knowledge of the world. Consequently, audio description must take this into account.

Translating cultural content represents a challenge in film translation because of spatio-temporal constraints, which may be addressed through strategies (Valero-Gisbert, 2021: 37) such as expansion, explication, or spatial dislocation¹².

5. Conclusion

This brief presentation allows us to draw a few considerations. First, there is the need to provide accessibility to cultural products so that everyone can freely decide, based on their own preferences, which cultural products they want to consume. Second, when applying audio description to audiovisual products in general, special attention must be given to the selection of words used for description and the syntax, as well as the length of the description. Although the description is limited by the film's duration, it cannot fill all the gaps the director intentionally left open. Additionally, the AD description speed must be balanced so that it does not demand excessive concentration and attention from the listener/receiver.

If the audio describer works with the production team, the editing can be done in a way that allows the most important details to be introduced, also emphasizing certain sounds so they contribute more clearly and carry greater semantic weight. Otherwise, in post-production, the lack of time or space may compromise the description of some important details, which sometimes forces the audio describer-translator to use other translation strategies such as compensation or spatial dislocation to provide essential information, or, when necessary, to omit it.

A culturally marked text, such as the selected film sequence, suggests conceiving audio description not only for the visually impaired but for all audiences. Experiencing visual content through AD, regardless of need, allows viewers to contrast what is perceived through the visual channel with their own knowledge. It opens a broader reflection on the power of images and fosters critical thinking.

References/ Reference List

- Aiello, G. (2020). Visual Semiotics: Key Concepts and New Directions. In L. Pauwels y D. Mannay (Eds.), *The SAGE Handbook of Visual Research Methods*, (pp. 367-380). London: SAGE.
- Ausburn, L. & Ausburn, F. (1978). Visual literacy: background, theory and practice, *PLET 15* (4), 291-297.
- Avgerinou, M. & Eriscon, J. (1997). A review of the concept of Visual Literacy. *British Journal of Educational Technology*, Vol 28 (4), 280-291
- Barthes, R. (1977 [1961]) 'The photographic message'. In *Image – Music – Text*. Trans. Stephen Heath. New York: Hill and Wang, (pp. 15–31).

¹² A curricular proposal will be presented in future work.

- Barthes, Roland (1977 [1964]) 'Rhetoric of the image'. In *Image – Music – Text*. Trans. Stephen Heath. New York: Hill and Wang, (pp. 32–51).
- Braun, S. (2011). Creating Coherence in Audio Description. *Meta: Journal des traducteurs / Meta: Translators' Journal*, 56(3), 645-662. Retrieved February 10, 2025, from <http://dx.doi.org/10.7202/1008338ar>
- Casetti, F. & De Chio, F. (1990). *Analisi del film*. Milano, Bompiani.
- Chion, M. (1994). *Audio-Vision: Sound on Screen*. In C. Gorbman, Trans. Columbia University Press.
- Cho, J.D., Jeong, J. Kim, J.H., & Lee, H. (2020). Sound Coding Color to Improve Artwork Appreciation by People with Visual Impairments. *Electronics* 2020, 9, 1981, 1-18.
- Convents, G. (2002). *I cattolici e il cinema*. In G.P. Brunetta, *Storia del cinema mondiale*, V, *Teoria, strumenti, memorie*, Torino, (pp. 485-517).
- Cosenza, D. (2023). *A Lacanian Reading of Anorexia*, Routledge, NYork.
- Debes, J. (1969). The loom of Visual Literacy: an overview. *Audiovisual Instruction* 14(8), 25-27.
- Diderot, D. *Lettera sui ciechi per l'utilità dei vedenti*, Italian and French, New Digital Press, Palermo, 2016.
- Dondis, D. A. (2006). *La sintaxis de la imagen. Introducción al alfabeto visual*. Barcelona: Gustavo Gili.
- Jakobson, R. (1959). On Linguistic Aspects of Translation. In R. Brower (ed.), *On Translation*. Cambridge Mass, Harvard University Press, (pp. 232-239).
- Kress G. & van Leeuwen, T. (1996). *Reading images. The grammar of visual design*. London: Routledge.
- Lacan, J. (1964). *The Four Fundamental Concepts of Psycho-Analysis*. Routledge: London & New York 2018.
- Maszerowska, A. (2012). Casting the light on cinema—how luminance and contrast patterns create meaning. *MonTI. Monografías de Traducción e Interpretación*, 4, 65-85. Retrieved February 10, 2025, from <https://doi.org/10.6035/MonTI.2012.4.3>
- Poulakis, N. & Tzamtzi, Z. D. (2023). Audiovisual Literacy and the Anthropology of Music. In J. Lee, W. Huang, X. Chen, F. Rodrigues, L. Okan, S. Beene, C. Huilcapi-Collantes (Eds.), *Connecting & Sharing: The Book of Selected Readings 2023*. International Visual Literacy Association. (pp. 140-149). Retrieved February 10, 2025, from <https://doi.org/10.52917/ivlatbsr.2023.020>
- Remael, A. et al. (2014). *Pictures Painted in Words. ADLAB Audio Description Guidelines*.

Valero-Gisbert, M.J. (2021). *La Audiodescripción: de la imagen a la palabra. Traducción intersemiótica de un texto multimodal*. Clueb, Bologna.

Valero-Gisbert, M.J. (2024). La traducción de la imagen en la audiodescripción cuando vehicula un contenido cultural. *HIKMA* 23(1), 315–336.

How to cite this article:

Valero Gisbert MJ and COSENZA D (2025) 'Words for Visual: Eliminating Barriers for Communication', *International Multidisciplinary Research Journal*, Volume:11; November 2025; Page 51-63. DOI: <https://doi.org/10.47722/imrj.2001.45>