Parallèles 37(2) October 2025

The audio description script through the lens of multimodality. A qualitative and quantitative analysis of the meaning codes in *Elite*

Alejandro Romero-Muñoz
Universitat Jaume I

Abstract

Multimodality has been a crucial concept in audiovisual translation (AVT) research that has allowed scholars to explore the audiovisual text. In this vein, researchers should move beyond the idea of AD being a homogenous notion to incorporate a multimodal view where the AD script is the result of different sign systems whose interaction conveys complex meanings. In this study, we first examine how the contents of AD scripts can be described using Chaume's (2004) proposal of the multimodal meaning codes. Secondly, we explore if this multimodal analysis can account for any difference between the English and Spanish versions. The interaction of multimodal information in an AD script fragment from the Netflix series *Elite* in both languages was addressed by resorting to a qualitative and quantitative analysis. Our results show that multimodality can describe the sign systems that comprise AD scripts. Moreover, the core of AD scripts in both languages consists of mainly visual information: movement, iconography, textual elements, and spatio-temporal changes. Therefore, we can state that *Elite*'s AD script in English and Spanish is constituted by four sources of visual information, which can help future researchers to study AD in a more precise way from different perspectives and methodologies.

Keywords

Audio description, multimodality, meaning codes, qualitative analysis, quantitative analysis

CONTACT

Romero-Muñoz Alejandro, alromero@uji.es, Universitat Jaume I, Avinguda Vicent Sos Baynat s/n, 12071 Castelló de la Plana, Spain

 DOI:10.17462/para.2025.02.03

1. Introduction

Audio description (AD) could be defined as an audiovisual translation (AVT) mode that aims to make audiovisual products and events accessible primarily for visually impaired and partially sighted users. AVT researchers have long used multimodal notions to describe the audiovisual text as the result of at least visual and acoustic information interacting to convey complex meanings. In this light, the motivation that lies behind this study is the belief that AD tends to be conceived as a homogenous theoretical notion where cultural references are transmitted, where differences between languages can be observed, where creative alternatives can be implemented, etc. However, since the AD script is also an audiovisual text, throughout this paper we will support the idea that a particular kind of multimodal analysis based on Chaume's (2004) meaning codes will allow researchers to consider AD scripts as the result of different sign systems so that future studies can focus on the precise sources of information conveyed in the scripts rather than on AD as a whole.

In this vein, the main research question that structures this paper is how different sign systems comprise the AD scripts and if the way these signs interrelate can be described by using multimodal meaning codes. Therefore, the main objective is to explore the way by which meaning codes can account for the multimodal contents of AD scripts. The resulting hypothesis is that signs from just some key meaning codes will be identified in the AD scripts. This assumption draws on Norris' (2004) notion of "modal configuration", which is the hierarchy of relevant sign systems or modes that arise when the interaction of modes produces particularly complex meanings. Additionally, this paper considers a second research question, which is whether the AD scripts in English and Spanish display differences in multimodal terms. Accordingly, the second objective is to delimitate which meaning codes constitute the multimodal components of AD in both languages. Finally, based on previous studies tackling the differences in AD among languages (Matamala & Rani, 2009), the second hypothesis is that the AD fragment in English and Spanish will consist of different multimodal contents.

2. Audiovisual translation and audio description

AVT is characterized by the interlingual and intralingual transfer of audiovisual texts, simultaneously conveying information through two communication channels (three if we consider the tactile channel, following Mejías-Climent, 2021) by which codes (composed of signs) convey meanings. According to Mangiron (2022) "AVT" is used as a hypernym that includes the translation and adaptation of multimedia, multimodal, and multisemiotic texts combining verbal and non-verbal signs transmitted through the acoustic and visual channels. Among the different AVT modes, dubbing and subtitling are probably the ones that have most attracted scholars' attention, although AVT includes many other modes: voice-over, AD, respeaking, subtitling for the deaf and hard of hearing (SDH), etc. Following other AVT classifications, Mangiron (2022) distinguishes between "traditional media" (films, series, documentaries, etc.) and "newer media" or "localization" (videogames, multimedia products, software, etc.).

Despite these proposals, which include AD and usually consider it as an accessible AVT mode, the relationship between AVT and accessibility is currently being reexamined. One of the approaches states that accessibility belongs to AVT, so dubbing or subtitling would be related to AD, with the particularity that the latter would be an accessible AVT mode. An early instance of this approach could be *The Arts and 504: a 504 Handbook for Accessible Arts Programming* (Greco & Jankowska, 2020), where references to "captions" and "verbal description" appear together. Another perspective reckons that accessibility should be considered as a different discipline: "media accessibility" or MA. In this regard, Greco (2018) supports the shift from

a particularist vision of accessibility (where accessibility would be aimed at specific groups of people) towards a universalist one (where accessibility would be aimed at every human being). Finally, a flexible viewpoint that envisions accessibility as part of AVT and, at the same time, allows the possibility of AVT being part of a wider concept of accessibility can be found in Matamala (2019), who identifies the following categories: linguistic accessibility, sensory accessibility, and cognitive accessibility. Therefore, accessibility would include prototypical modes such as SDH or AD (sensory accessibility), but also dubbing or subtitling (linguistic accessibility). If we conceive accessibility as part of AVT, AVT would include prototypical modes, like dubbing and subtitling, but also other accessible modes, like AD and SDH.

Accordingly, AD could be defined as an AVT or MA mode that aims to make audiovisual products and events accessible for visually impaired and partially sighted users. Vercauteren states that AD translates "the visual and aural elements that [people with sight loss] do not have access to into a verbal commentary" (2022, p. 78). This verbal commentary should be embedded in silent moments so as not to overlap with relevant aural elements. According to Mazur (2020, p. 228), AD primary users are visually impaired or partially sighted people, but secondary users with no vision impairment can also be found (elderly people, people with cognitive difficulties, users employing AD with learning purposes, etc.). Researchers have resorted to Jakobson's (1959) concept of "intersemiotic translation" to state that AD is a translation from a non-verbal sign system (the image and some sounds) into a verbal sign system (the AD script), although AD can also be interlingual when AD scripts are translated (Jankowska, 2015; Romero-Muñoz, 2024). Beyond this, Jankowska (2015) links AD to AVT concluding that AD meets three AVT requirements: its purpose is to make national or foreign audiovisual products available to people that would not understand them without that help; both the ST and the TT consist of audiovisual or multimodal texts, where several meaning codes are transferred by means of the visual and acoustic channels; and AD shares important characteristics with other AVT modes, like space and time restrictions.

3. A multimodal view of audio description

In many definitions of multimodality, the notions of semiotic modes and the way they interact seem to be what characterize the discipline. Kress & Van Leeuwen (2001, p. 20) refer to multimodality as: "the use of several semiotic modes in the design of a semiotic product or event, together with the particular way in which these modes are combined". Semiotic modes can be described as "semiotic resources", whose meaning is "culturally made, socially agreed and socially and culturally specific" (Kress 2014, p. 60). Regarding the interrelation of modes, it is this interaction that creates meaning O'Halloran, K.L.E. & Tan, 2015); hence multimodality focuses on how semiotic modes combine to constitute the multimodal text (Taylor, 2020).

Following Villanueva-Jordán (2024), four concepts are involved in the creation of meaning in multimodal texts: the resource integration principle, multiplying meaning, modal density, and modal configuration. On the one hand, the resource integration principle was proposed by Baldry & Thibault (2006), and it deals with how modes or semiotic resources relate to each other. On the other hand, Lemke states that meanings stemming from different semiotic modes influence meanings from other semiotic modes, "thus *multiplying* the set of possible meanings that can be made" (1998, p. 92). Finally, Norris (2004) uses the notion of "modal density" as the intensity and complexity achieved by means of the interaction of different modes, and because of this density some modes might become more relevant than others, a hierarchy that Norris (2004) labels as "modal configuration".

After overviewing the basic concepts of multimodality, it seems safe to support Kourdis' (2022) view that semiotics constitutes the basis of multimodality and multimodality is in turn one

of the theoretical frameworks used by AVT scholars. The way semiotics, multimodality, and AVT connect can be seen in many definitions of the audiovisual text. Díaz Cintas (2020) states that audiovisual texts are semiotic composites in which codes merge to create meaning, while Chaume (2020, p. 108) defines these texts as "semiotic constructs where meaning is produced by combining different signs, encoded in various codes and transmitted through at least two channels of communication: acoustic and visual". Many scholars have studied the nature of the audiovisual text, such as Zabalbeascoa (2001), whose classification highlights two parameters that specifically affect AVT: channels of communication (audio and visual) and sign codes (verbal and non-verbal). Therefore, elements comprising the audiovisual text should belong to one of these four dimensions, which reveals its semiotic complexity and multimodal essence. Chaume (2004; 2012) has explored the characteristics of the audiovisual text by proposing an analysis based on the classification of eleven meaning codes transmitted either by the acoustic or the visual channel that specifically affect AVT. These codes are constituted by signs and their meaning is conventionalized by culture, which relates to the traditional multimodal notion of "mode" (Villanueva-Jordán, 2024).

If we consider the meaning codes transmitted through the acoustic channel, one of the most important codes in some AVT modes is the linguistic one, and it refers to any language information that can appear in dialogues, monologues, a narrator's voice, etc. The paralinguistic code has to do with aural nonverbal information, such as laughter, clicks, whispering sounds, voice volume, etc. Regarding the musical code, here we could include information related to the soundtrack and songs. The special effects code alludes to sound elements not uttered by characters. Finally, the sound position code deals with the origin of characters' voices.

Moving on to the codes transmitted through the visual channel, the iconographic code is composed of indexes, icons, and symbols. Conversely, the photographic code alludes to information related to perspective, light, and color. As for the mobility code, we can distinguish between proxemic signs (the characters' distance from each other and their distance from the camera), kinesic signs (body movements and gestures), and mouth articulation signs. The shot code refers to the use of camera movements and angles, as well as the information they convey. The graphic code deals with written language that might appear on screen, such as titles, intertitles, subtitles, credits, etc. Finally, the editing code has to do with the film transition marks that organize the film in shots, sequences and so forth (fade outs, iris wipes, cuts to, etc.).

As previously clarified, there is a connection among semiotics, multimodality and AVT (which includes AD). The semiotic nature of AD can be hinted in the way many researchers (Matamala, 2019; Taylor, 2020; Fresno, 2022; and Taylor & Perego, 2022) use Jakobson's (1959) concept of intersemiotic translation. However, multimodality and AD can also be associated: "multimodality can be described as a defining feature of AVT, but in the cases of SDH and AD this is all the clearer" (Taylor 2020, p. 84). In fact, some studies have already explored AD from a multimodal perspective, such as Hirvonen & Tiittula (2010), Álvarez de Morales (2011), Braun (2011), the TRACCE corpus (Jiménez & Seibel, 2012), Jiménez Hurtado & Soler Gallego (2013), Chica Núñez (2015), Carlucci & Seibel (2016), Randaccio (2018), Reviers (2018), Matamala (2019), Remael & Reviers (2019), Taylor (2019; 2020), or Holsanova (2020), among others.

Hirvonen & Tiittula (2010) conceive texts as multimodal sign systems, and these authors present a method to analyze multimodal texts applied to AD. Álvarez de Morales (2011) tackles the AMATRA project to present AD as a new type of discourse related to Quintilian's rhetoric. In Braun's (2011) study, the author explores how the coherence of a source multimodal text (a film) is recreated in a target multimodal text (AD). The TRACCE corpus contains audio described films labelled using three semiotic dimensions: narratology, cinematography, and grammar (Jiménez & Seibel, 2012). Jiménez Hurtado & Soler Gallego (2013) explain how to

analyze AD by using the multimodal annotation software Taggetti applied to the TRACCE corpus. Chica Núñez (2015) focuses on color and movement in AD to study the multimodal perception, which he identifies as the translation process that occurs in AD when the audio describer perceives audiovisual elements and produces a functionally equivalent text. Carlucci & Seibel (2016) analyze the DESAM project, which seeks to design teaching strategies to create a multimodal space for Translation and Interpreting students. Randaccio (2018) connects the changes in Museum Studies during the 1980s and 1990s with the development of museum AD, which is conceived as a multimodal and multisensory translation. Reviers (2018) employs a corpus-based multimodal study to describe the linguistic features of AD scripts and the role they play in the communicative function. Matamala's (2019) VIW (Visual Into Words) project creates a multimodal AD corpus of AD in English, Spanish, and Catalan annotated by means of the multimodal corpus analysis tool ELAN. Remael & Reviers (2019) illustrate how multimodal cohesion is maintained or (re)created in an accessible film clip with AD and SDH using Tseng's (2013) model. Taylor (2019; 2020), combines theoretical approaches like cognitive linguistics, systemic-functional linguistics, and discourse analysis to move beyond the image or word level and provide AD for museums with other senses (music, sounds, smell, and taste). Finally, Holsanova (2020) states that scientific texts are multimodal, and they display images that can be challenging for readers, that is why she proposes AD as a tool to help readers to attain higher multimodal literacy.

Even if these studies have enriched the development of research on AD, they use very different theoretical approximations (Quintilian's rhetoric, TRACCE's semiotic dimensions, Tseng's model, etc.) or analysis tools (ELAN and Taggetti) that hinder the possibility of associating their contributions to the field. Moreover, it is our view that on many occasions research on AD tends to conceive this AVT or MA mode as a homogeneous notion instead of a multimodal text where different sign systems or modes combine. Consequently, despite being very valuable contributions, we should perhaps move beyond the idea of cultural references (Jankowska, 2022), interlinguistic differences (Matamala & Rani, 2009) or creative alternatives (Holsanova, 2016; Bardini, 2020; Soler Gallego & Luque Colmenero, 2023) being studied in a general and abstract notion of AD. Since subtitling research does not focus on the abstract concept of subtitles, but rather on line breaks, colors, number of characters, or subtitles position, for instance, we should discuss what precise sources of information tend to convey cultural references, what sign systems can be more creative, what codes entail bigger differences among languages, etc.

As it will be argued throughout the paper, this operationalization of AD contents can be achieved through multimodal meaning codes. Accordingly, in the following section we will provide a methodological proposal to incorporate meaning codes in the analysis of AD to fully account for the sign systems that interact to create meaning in an AD script in English and Spanish.

4. Methods and analysis

Throughout this research, we employed an exploratory methodology based on a case study, namely the Netflix series *Elite* in English and Spanish. This methodological choice seeks to determine the possibility of studying from a multimodal perspective how different sign systems constitute the AD scripts and the way these signs interrelate in both languages. It is our view that an analysis based on meaning codes can operationalize the AD scripts' contents in multimodal terms, which would allow future studies to focus on the basic components that constitute AD rather than on an abstract and general notion of AD.

As for the material used in the analysis, *Elite* is the result of three levels of selection criteria,

in each of which a corpus was compiled to eventually select this series (see Romero-Muñoz, 2023 for a thorough explanation of every step). As the basis of the subsequent analysis, the first selection level needed to be rigorous and supported by objective principles. Therefore, out of the considerable amount of available material, the most appropriate audiovisual products were chosen based on availability, production, linguistic, and temporal filters. The application of these criteria resulted in corpus 0, a database consisting of 40 Netflix series from 2022 containing English and Spanish AD. Analyzing the episodes of 40 series seemed to be an excessive task, so drawing on corpus 0, we configured a further filtered database called "corpus 1", whose aim was to focus on an appropriate number of series with common features to make the analysis more effective. To compile corpus 1, we established another three selection criteria (series origin, date, and genre), which resulted in one drama and thriller Spanish Netflix series released in 2022: *Elite.* Even after having created corpus 1, there was too much material, so it was necessary to establish some sort of selection principle to narrow down which episodes and scenes to analyze (corpus 2). Consequently, we selected approximately the first five minutes of the first episode from *Elite*'s latest season, starting from the very beginning up to an appropriate scene or sequence change.

Once we had selected the ideal fragment for the analysis, we transcribed the English and Spanish AD scripts. To do so, we opted for Transkriptor, a software program that transcribes audio into text in several languages, and then a thorough revision was carried out to ensure proper quality standards. We did not limit the transcriptions to the AD fragments, but we also included dialogues so that the interaction between AD and other semiotic information could be more easily examined if needed.

Having transcribed the dialogues and AD fragments, we then applied the multimodal analysis based on Chaume's (2004) classification of the eleven meaning codes. As stated before, some of these codes are transmitted through the visual channel and others through the acoustic channel. Taking into account the characteristics and needs of the target audience of AD, we incorporated all the codes in the analysis to fully characterize any possible multimodal component of the script. To associate the sign systems interrelating in AD with meaning codes, we applied the principles of qualitative content analysis (QCA). According to Schreier (2012), QCA describes the meaning of qualitative material by classifying instances of certain categories in that material. These categories are labelled by means of codes, which in the QCA tradition means words or short phrases that we associate as attributes of the phenomenon that we are studying. In QCA, codes depend on a codebook, which is "a set of codes, definitions, and examples used as a guide to help analyze interview data […] they provide a formalized operationalization of the codes" (DeCuir-Gunby, Marshall, & McCulloch 2011, p. 138). Thus, Chaume's meaning codes were adapted to the characteristics of AD, examples of how signs from every code could manifest in an AD script were provided and the resulting combination constituted the codebook for the analysis (see Romero-Muñoz, 2023 for a full account of the codebook structure). In order to easily analyze the AD scripts' contents from a multimodal perspective, we represented each code by means of a symbol consisting of its first letter(s). In this way, (M) stands for a sign from the musical code in the analysis. Therefore, the AD script was analyzed by cataloguing its signs as instances belonging to any of the eleven meaning codes. With the aim of further describing the multimodal composition of AD scripts, it seemed appropriate to corroborate to what extent acoustic and visual codes appeared in AD. Seeking to easily verify a possible balance (or lack of balance), visual codes were colored in blue and acoustic codes in red (Table 1 and Table 2). Given the exploratory nature of this case study, where the main interest focused on the applicability of meaning codes on AD scripts, the coding phase was conducted by one researcher, who resorted to the codebook to systematise

how signs would fit in each category to minimize the subjectivity involved in the individual coding. Future descriptive studies with a bigger corpus should incorporate more researchers to annotate the textual material, with internal validations processes among coders, and constant comparisons of the results.

***Elite* (E1S4 – "The New Order")**
[54:33] A red letter N (I) unfolds (MB) into a spectrum of colors (PH). [54:29] Words (I) appear (MB): "A Netflix original series. A Zeta Studios Production" (G). In the Club del Lago (ED) Guzmán breathes heavily (P). **POLICE OFFICER**: Guzmán. Guzmán. I need to ask you a few questions. What happened? [54:12] Through the windows behind him (SHT) fireworks irrupt (SE) in the night sky (ED). **POLICE OFFICER**: Look, I need you to tell me what you saw here tonight. What happened? […] [53:30] He glances (MB) towards the detective (I). Two red mirror image "E" (I) slide out in opposite directions (MB) from the center of a black background (ED). They flip around (MB) as more letters appear (MB). Title: "Elite" (G). "Created by Carlos Montero and Darío Madrona" (G). In the Las Encinas campus pool (ED) Guzmán stands on the starting block (MB). He touches his toes (MB) then tilts his head to either side (MB). **TEACHER**: Boys, lanes one and two. Girls, three, four, and five. **ARI**: Is there any particular reason for this? **TEACHER**: For what? **ARI**: For dividing us by gender and not by skill level. The best swimmers could be in lane one and the worst in lane five. **TEACHER**: Just split up like I told you to. [52:47] The boys wear navy-blue trunks (I) and the girls are in red one-piece suits (I). The girl switches places with a boy (MB). **TEACHER**: On your marks… [52:35] As the race begins (MB), Samuel takes interest from the sidelines (MB). Guzmán is in the lane closest to him and reaches the wall first (MB), edging the girl in the middle lane (MB). Samuel smirks in approval (MB) as he watches the girl climb out (MB). On the deck (ED), Guzmán removes (MB) his cap (I) and goggles (I), then shakes water off his face (MB). Removing (MB) her own cap (I) and goggles (I), the girl runs her hand (MB) through her short dark hair (I). Guzmán beams at her (MB). With a thin smirk (MB), she brushes past (MB). He keeps his eyes on her (MB). She turns (MB) and makes eye contact (MB), then continues on (MB). Passing Sam (MB), she grins (MB). Now (ED) on a covered campus walkway (ED), Ander walks arm in arm (MB) with his mother (I), Azucena, the school principal (I). He kisses her cheek (MB). On the ceiling (SHT) is the Las Encinas' crest (I), featuring a large "E" (I) and two leaves (I).

Police Officer: Look, I need you to tell me what you saw here tonight. What happened? [54:02] A detective (I) stands before him (MB). He has handcuffs on his belt (I). Guzmán blinks (MB). Now (ED), a girl in a flowing blue dress (I) floats face-up (MB) in the lake (ED), close to the club's stadium (ED). Fireworks reflect off the rippling water (SE). Investigators (I) process the scene (MB). A rescue boat is anchored nearby (MB).

Table 1. Excerpt from the English AD analysis

Considering the English excerpt appearing in Table 1, almost every sign is blue, so resorting to the resource integration principle (Baldry & Thibault, 2006), we see that it is mostly visual information that interrelates in this AD script. References to how characters or things move ("flip", "tilt", "remove", "turn", "process", etc.) seem to be very frequent. After that, references to iconography ("cap", "goggles", or "Las Encinas' crest"), graphic elements ("A Netflix original series. A Zeta Studios Production"), or changes in time and space ("in the Las Encinas campus pool", "now") can be found. Therefore, probably due to the space and time restrictions for AD to convey information, the modal density (Norris, 2004) seems to be focused on information related to movement, iconography, graphic elements, and spatio-temporal changes. However, these sign systems are not equally important, since a modal configuration (Norris, 2004) is noticeable: movement and iconography are more frequent (so hierarchically more important) in the AD script, whereas the presence of graphic elements and spatio-temporal changes is more reduced.

As for the way signs have been coded, "a red letter N" is tagged as an iconographic sign, since it is a conventional symbol that means "Netflix". "Unfolds" is considered a mobility sign because the AD conveys information about the movement of this letter on screen. "Into a spectrum of colors" is a photographic sign because it alludes to color information. "Words" refers to the fact that a text appears on screen, while "A Netflix original series. A Zeta Studios Production" conveys the content of that text, so they are two different signs belonging to the graphic code. "In the Club del Lago" frames the scene in a particular location that contributes to the spatial progression of the fragment, so it is labelled as an editing sign. Finally, "Guzmán breathes heavily" is a paralinguistic sign that conveys acoustic non-verbal information uttered by a character. It must be noted that references to sight (such as "he glances") have been interpreted as movement signs, because they usually use the characters' facial expression to infer sight, even when the image depicts someone turning their head or frowning, for instance.

***Elite* (E1S4 – "The New Order")**
[54:33] Una letra N roja (I) se despliega (MB) y adopta varios colores (PH). [54:29] «Una serie original de Netflix» (G). «Una producción de Zeta Studios» (G). En el Club del Lago (ED). **Policía**: Guzmán. Guzmán. Tengo que hacerte unas preguntas. ¿Qué ha pasado? Necesito que me cuentes qué has visto esta noche. ¿Qué ha ocurrido? [54:02] Un policía (I) se para delante de Guzmán (MB). El cuerpo de una chica morena de pelo corto (I) con un vestido blanco (I) flota boca arriba (MB) en el lago (ED). [53:47] Sobre el lago (ED) los fuegos artificiales iluminan (SE) el cielo nocturno (ED).

POLICÍA: Guzmán [53:41] Sentado en una silla (MB), Guzmán mira al suelo (MB). Nervioso (MB), levanta la cabeza (MB). **POLICÍA**: Tranquilo. Todo va a salir bien. [53:30] Guzmán mira fijamente (MB) al policía (I). Del centro de un fondo negro (ED) surgen letras rojas (MB): «Élite» (G). «Creada por Carlos Montero y Darío Madrona» (G). [53:11] De día (ED), en la piscina de Las Encinas (ED) Guzmán está parado (MB) sobre un trampolín (I), lleva puesto un bañador (I), gorro (I) y gafas de piscina (I).

Table 2. Excerpt from the Spanish AD analysis

Regarding the Spanish fragment appearing in Table 2, again virtually every sign is blue, which aligns with the previous tendency, since visual information seems to be prioritized in the Spanish AD script too. References to how characters or things move ("se despliega", "sentado", "levanta", "se para", etc.) are very frequent in Spanish as well. Similarly to the previous case, allusions to iconography ("policía", "bañador", or "gafas de piscina"), graphic elements ("Una serie original de Netflix"), or spatio-temporal changes ("En el Club del Lago", "el cielo nocturno") are found. The modal density (Norris, 2004) focuses again on information related to movement, iconography, graphic elements, and spatio-temporal changes. Following the English AD tendency, the Spanish script shares the same modal configuration (Norris, 2004): movement and iconography are more frequent than graphic elements and spatio-temporal changes. Even if this might seem to contradict other studies on the interlinguistic differences among AD, such as Matamala & Rami (2009), they point at differences in the amount of information provided, not in the type of information, that is why research on the sources of information conveyed by AD is needed.

5. Results

From a qualitative perspective, and according to Baldry & Thibault's (2006) resource integration principle, our AD script seems to be composed mainly of interrelating visual elements in both languages. After that, the scripts' modal density (Norris, 2004) narrows down to sign systems related to movement, iconography, graphic elements, and time or space changes. However, these four sources of information display a certain hierarchy or modal configuration (Norris, 2004), since refences to movement or iconography are more frequent than graphic elements and time or space changes. Beyond that, English and Spanish do not seem to have significant differences when it comes to the sign systems that interact in the AD scripts.

Moreover, the qualitative data can be contrasted with quantitative data obtained by measuring how frequently signs from each code appear in the AD scripts. A closer examination allows us to delimitate the exact frequency with which signs from every code have appeared throughout the AD scripts in both languages. Focusing first on the overall results in English and Spanish, only 1 sign from the paralinguistic code has been found (0.61 % of the AD scripts), 3 signs from the special effects code (1.83 %), 85 signs from the mobility code (51.83 %), 38 signs from the iconographic code (23.17 %), 26 signs from the editing code (15.85 %), 7 signs from the graphic code (4.27 %), 2 signs from the photographic code (1.22 %), and 2 signs from the shot code (1.22 %). On the other hand, if we consider the data from the English AD, in the acoustic channel there was 1 sign from the paralinguistic code (1.02 % of the AD script in English), 2 signs from the special effects codes (2.04 %), 50 signs from the mobility code (51.02 %), 25 signs from the iconographic code (25.51 %), 14 signs from the editing code (14.29 %), 3 signs from the graphic

code (3.06 %), 1 sign from the photographic code (1.02%), and 2 signs from the shot code (2.04%). Finally, if we focus on the results from the Spanish AD, there was 1 sign from the special effects code (1.51 % of the AD script in Spanish), 35 signs from the mobility code (53.03 %), 13 signs from the iconographic code (19.71 %), 12 signs from the editing code (18.18 %), 4 signs from the graphic code (6.06 %), and 1 sign from the photographic code (1.51 %).

Consequently, the quantitative data obtained allows the description of the multimodal configuration of the AD script in English and Spanish as represented in Figure 1. Similarly, the multimodal configuration of the AD script in English is depicted in Figure 2, while the Spanish one can be found in Figure 3. As we can see, these tendencies seem to align with the qualitative data previously explained. Be that as it may, this multimodal analysis based on meaning codes has enabled the precise delimitation of the semiotic elements that configurate and interact in the English and Spanish AD excerpts.

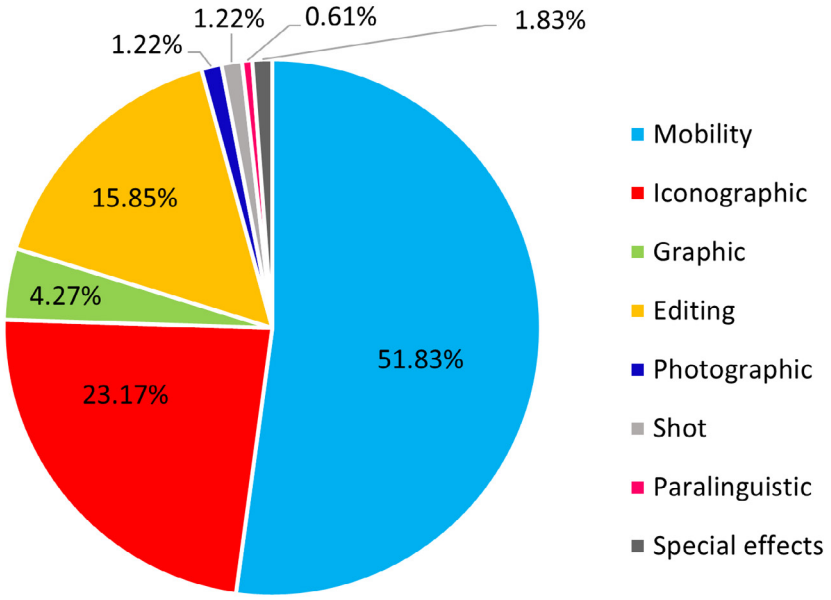

Figure 1. Multimodal configuration of the AD script in English and Spanish

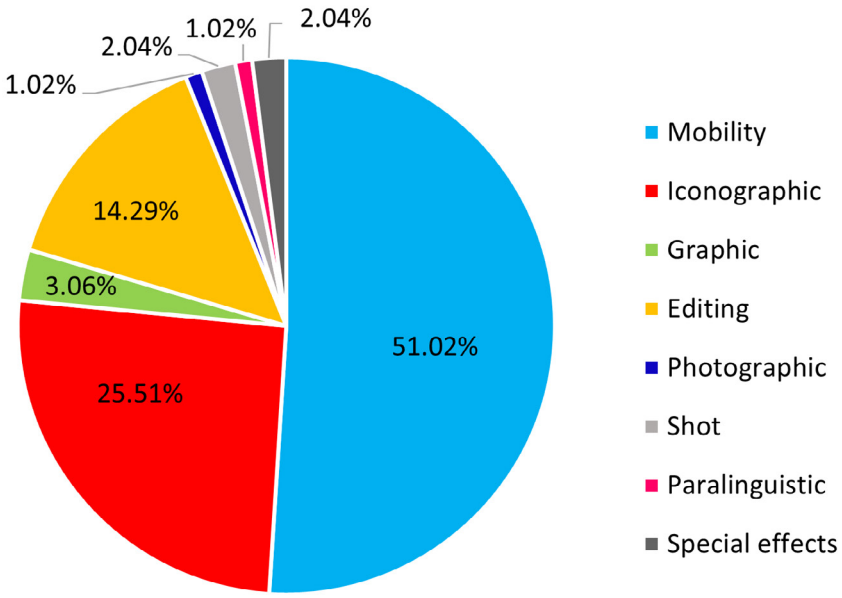

Figure 2. Multimodal configuration of the AD script in English

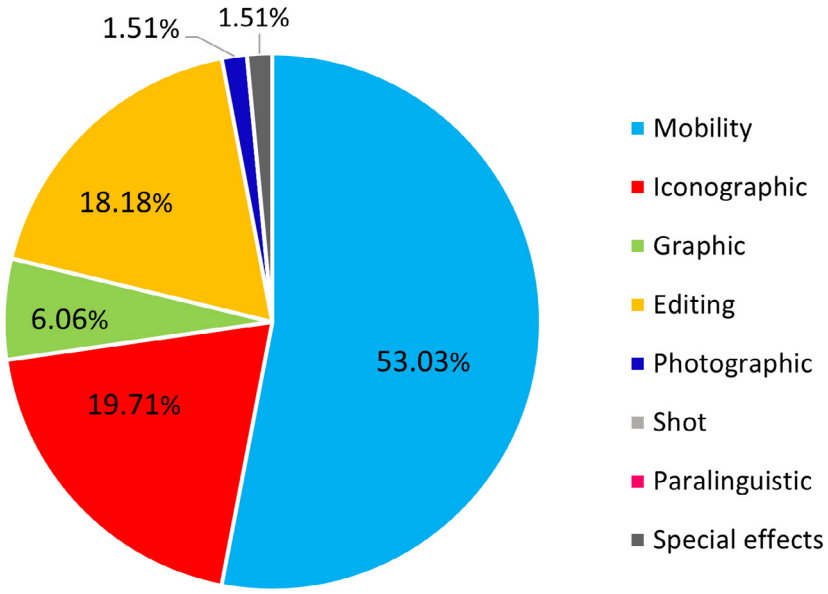

Figure 3. Multimodal configuration of the AD script in Spanish

6. Discussion and conclusions

From a qualitative perspective, and according to Baldry & Thibault's (2006) resource integration principle, our AD script seems to be composed mainly of interrelating visual elements in both AD. As we have stated throughout this paper, our main research question is how different sign systems comprise the AD scripts and if the way these signs interrelate can be described by using multimodal meaning codes. Considering this research question, one of the most remarkable conclusions resulting from the qualitative and quantitative data in both languages

is the lack of balance in terms of channels. Since acoustic codes are so rare in our fragments, we can safely say that our AD scripts are essentially composed of visual sign systems. Even if this fact should not be a surprise taking into account the AD audience, it is important to verify from a qualitative and quantitative perspective that AD conveys the information that users need, at least in sensory terms. Moreover, given the fact that aural elements comprise a fraction of the semiotic information conveyed in the AD script, the need for a fully articulated multimodal approach based on all meaning codes seems reasonable. Bearing in mind that virtually the entire multimodal configuration of the AD script is constituted by visual elements, further regularities can be observed, which helps us achieve the main objective of exploring the way by which meaning codes can account for the multimodal contents of AD scripts. References to movement and iconography encompass approximately three quarters of the information provided in the AD script. After that, graphic and editing codes virtually constitute the remaining quarter of the AD script. Finally, photographic, shot, paralinguistic, and special effects codes are particularly rare, whereas the rest of the codes do not appear in our fragments. Therefore, it can be stated that the core of our AD excerpts revolves around four codes (mobility, iconographic, editing, and graphic), which establish nearly 95% of the AD scripts. This confirms our first hypothesis, because signs from just four meaning codes constitute the multimodal configuration of our AD scripts to convey complex meanings.

Furthermore, a second research question also considered whether the AD scripts in English and Spanish displayed differences in multimodal terms. As far as the linguistic dimension is concerned, most of the tendencies can be observed in Spanish and English without important differences. In both languages visual contents prevail over aural information. Similarly, the four main codes of the AD script remain the same in English and Spanish, which helps us achieve the second objective of delimitating which meaning codes constitute the multimodal components of AD in both languages. Nevertheless, English seems to describe more iconographic information, whereas Spanish focuses more frequently on graphic and editing elements. Finally, perhaps the most evident difference is that English seems to provide more information than Spanish, since the first has described 98 signs, whereas the latter included 66, which aligns with Matamala & Rani's (2009) findings on interlinguistic differences among AD. It is our view that these differences are only partially true, since at least in our case the main difference between English and Spanish AD is attributed to the amount of information provided, but the sources of information used to convey meaning are the same. Since this analysis has shed some light on the semiotic components of the AD script, very similar in both languages, the second hypothesis is not completely confirmed.

This leads to the following question: why do these meaning codes constitute the AD scripts? We might venture a tentative response that should be further explored in the future. Firstly, the reason why mobility abounds might be attributed to the characteristics of drama and thriller series, where action is usually involved. Secondly, the inclusion of iconographic elements is crucial in AD, since symbols, indexes or icons constitute a considerable source of semiotic information in any scene. Moreover, graphic signs are numerous in our excerpts, likely due to the first-five-minute selection principle applied to the episode, where initial credits are expected. Finally, editing signs are frequent in the excerpts because they provide valuable information about place and time progressions that help AD users to better understand the plot, whose development might be subject to particularly quick temporal and physical shifts in thriller and drama series.

One last remark worth reflecting on is the research potential of multimodality applied to AD. What these results prove is that a key advantage of a multimodal analysis based on meaning codes lies in its ability to isolate very specific pieces of information encapsulated in signs. In other

words, multimodal meaning codes allow the operationalization of the semiotic information conveyed in the AD scripts, which can provide researchers with a more precise view of AD as a multimodal text where different sources of information interact. Instead of envisioning AD as a general abstract notion, this operationalization could lead to other research methodologies or approaches tackling any AD components: by means of which codes are cultural references usually rendered in AD? What is the multimodal configuration of an AD corpus? Do users accept the presence of more acoustic elements?

Finally, the data obtained in this exploratory case study should not be extrapolated to any other material beyond our excerpts. They should be taken as a first methodological step that needs to be further explored by integrating other methodologies, such as descriptive corpus-based or experimental reception studies.

7. Acknowledgements

This research has been developed with the financial support of the predoctoral contract FPI-UJI/2021 (Programa propi UJI) granted by the Universitat Jaume I.

8. References

Álvarez de Morales, C. (2011). Audio Description, Rhetoric and Multimodality. *Journal of Language Teaching and Research*, *2*(5), 949-956. https://doi.org/10.4304/jltr.2.5.949-956

Baldry, A. & Thibault, P. (2006). *Multimodal Transcription and Text Analysis. A Multimedia Toolkit and Coursebooks with Associated On-line Course*. London: Equinox.

Bardini, F. (2020). *La transposició del llenguatge cinematogràfic en l'audiodescripció i l'experiència fílmica de les persones amb discapacitat visual.* Doctoral dissertation, Universitat de Vic*.* Tesis Doctorals en Xarxa (TXD) Repository. https://www.tesisenred.net/handle/10803/669901#page=1

Braun, S. (2011). Creating Coherence in Audio Description. *Meta: Translators' Journal, 56*(3), 645-662. https://doi.org/10.7202/1008338ar

Carlucci, L. & Seibel, C. (2016). Universidad, accesibilidad y nuevas tecnologías. Valoración de una experiencia de innovación docente en la traducción especializada. *Revista Didáctica, Innovación y Multimedia*, *33*, 1-16.

Chaume, F. (2004). *Cine y traducción*. Madrid: Cátedra.

Chaume, F. (2012). *Audiovisual Translation: Dubbing*. Manchester: St. Jerome Publishing.

Chaume, F. (2020). Dubbing. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 103-132). London*:* Palgrave Macmillan.

Chica Núñez, A. J. (2015). Multimodality and multi-sensoriality as basis for access to knowledge in translation. The case of audio description of colour and movement? *Procedia Social and Behavioral Sciences, 212,* 210-217. https://doi.org/10.1016/j.sbspro.2015.11.335

DeCuir-Gunby, J., Marshall, P. & McCulloch, A. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. *Field Methods, 23*(2), 136-155.

Díaz Cintas, J. (2020). An Excursus on Audiovisual Translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 11-32). London*:* Palgrave Macmillan.

Fresno, N. (2022). Research in audio description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 312-327). New York: Routledge.

Greco, G. M. (2018). The Nature of Accessibility Studies. *Journal of Audiovisual Translation, 1*, 205-232. https://doi.org/10.47476/jat.v1i1.51

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). London*:* Palgrave Macmillan.

Hirvonen, M. & Tiittula, L. (2010). A method for analysing multimodal research material. Audio description in focus. *Suomen kääntäjien ja tulkkien, 4*.

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.

Holsanova, J. (2020). Uncovering scientific and multimodal literacy through audio description. *Journal of Visual Literacy*, *39*(3-4), 132-148. https://doi.org/10.1080/1051144X.2020.1826219

Jankowska, A. (2015). *Translating audio description scripts. Translation as a new strategy of creating audio description*. Frankfurt: Peter Lang.

Jankowska, A. (2022). Audio description and culture-specific elements. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 107-126). New York: Routledge.

Jakobson, R. (1959). On Linguistic Aspects of Translation. In R. Arthur Brower (Ed.), *On Translation* (pp. 232-239). Cambridge: Harvard University Press.

Jiménez, C. & Sibel, C. (2012). Multisemiotic and Multimodal Corpus Analysis in Audio Description: TRACCE. In A. Remael, P. Orero & M. Carroll (Eds.), *Audiovisual Translation and Media Accessibility at the Crossroads: Media for All* 3 (pp. 409-425). Rodopi. https://doi.org/10.1163/9789401207812_022

Jiménez Hurtado, C. & Soler Gallego, S. (2013). Multimodality, translation and accessibility: A corpus-based study of audio description. *Perspectives. Studies in Translatology, 21*(4), 577-594. https://doi.org/10.1080/0907676X.2013.831921

Kourdis, E. (2022). Semiotics. In F. Zanettin & C. Rundle (Eds.), *The Routledge Handbook of Translation and Methodology* (pp. 139-154). London*:* Routledge.

Kress, G. (2014). What Is Mode? In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 60-75). London: Routledge.

Kress, G. & Van Leeuwen, T. (2001). *Multimodal Discourse: the Modes and Media of Contemporary Communication*. London: Arnold.

Lemke, J. (1998). Multiplying meaning. Visual and verbal semiotics in scientific text. In J. Martin & R. Veel (Eds.), *Reading Science: Critical and Functional Perspectives on Discourses of Science* (pp. 87-113). London*:* Routledge.

Mangiron, C. (2022). Audiovisual Translation and Multimedia Localisation. In F. Zanettin & C. Rundle (Eds.), *The Routledge Handbook of Translation and Methodology* (pp. 410-424). London*:* Routledge.

Matamala, A. (2019). The VIW corpus: multimodal corpus linguistics for audio description analysis. *RESLA. Revista Española de Lingüística Aplicada 32*(2), 515-542. https://doi.org/10.1075/resla.17001.mat

Matamala, A. & Rami, N. (2009). Análisis comparativo de la audiodescripción española y alemana de *Good-bye, Lenin*. *Hermeus*, *11*, 249-266.

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. 227-2247). London*:* Palgrave Macmillan.

Mejías-Climent, L. (2021). *Enhancing Video Game Localization Through Dubbing*. *London:* Palgrave Macmillan.

Norris, S. (2004). *Analyzing multimodal Interaction: A Methodological framework*. Routledge.

O'Halloran, K., K. L. E., M. & Tan, S. (2015). Multimodal Semiosis and Semiotics. In J. Webster (Ed.), *The Bloomsbury Companion to M. A. K. Halliday* (pp. 386-411). London: Bloomsbury.

Randaccio, M. (2018). Museum Audio Description: Multimodal and 'Multisensory' Translation: A Case Study from the British Museum. *Linguistics and Literature Studies*, *6*(6), 285-297. https://doi.org/10.13189/lls.2018.060604

Remael, A. & Reviers, N. (2019). Multimodality and audiovisual translation. Cohesion in accessible films. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 260-280). London: Routledge.

Reviers, N. (2018). *Audio Description in Dutch: A Corpus-Based Study into the Linguistic Features of a New, Multimodal Text Type.* Doctoral dissertation, University of Antwerp.

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. (2024). La traducción de audiodescripciones. Situación en las plataformas de *streaming* y propuestas didácticas para la L1 y L2. In R. Martínez, C. Pulido & B. Mateu (Eds.), *Nuevas investigaciones educativas para definir la enseñanza y el aprendizaje* (pp. 721-728). Octaedro.

Soler Gallego, S. & Olalla Luque Colmenero, M. (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

Taylor, C. (2019). Audio description: A multimodal practice in expansion. In J. Wildfeuer, J. Pflaeging, J. Bateman, O. Seizov & C.I. Tseng (Eds.), *Multimodality. Disciplinary thoughts and the challenge of diversity* (pp. 195-218). Berlin: De Gruyter Mouton.

Taylor, C. (2020). Multimodality and Intersemiotic Translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 83-99). London*:* Palgrave Macmillan.

Taylor, C. & Perego, E. (2022). Museum audio description: the role of ADLAB PRO. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 38-54). New York: Routledge.

Tseng, C.I. (2013). *Cohesion in Film. Tracking Film Elements*. Basingstoke: Palgrave Macmillan.

Vercauteren, G. (2022). Narratology and/in Audio Description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 78-92). New York: Routledge.

Villanueva-Jordán, I. (2024). *Traducción audiovisual y teleficción queer. Teoría y metodología*. Lima: Editorial Universidad Peruana de Ciencias Aplicadas.

Zabalbeascoa, P. (2001). El texto audiovisual: factores semióticos y traducción. In J. Sanderson (Ed.), *¡Doble o Nada!* (pp. 113-126). Alicante*:* Publicacions de la Universitat de Alacant.