

## Multimodal Analysis as a Way to Operationalise Objectivity in Audio Description: A Corpus-based Study of Spanish Series on Netflix

 Alejandro Romero-Muñoz 

Universitat Jaume I

Citation: 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>

**Editor(s):** E. Schaeffer-Lacroix, N. Reviere & E. Di Giovanni

**Received:** November 1, 2023

**Accepted:** June 29, 2023

**Published:** December 27, 2023

**Copyright:** ©2023 Romero-Muñoz. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/). This allows for unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Acknowledgements

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

### Abstract

Objectivity is a constant recommendation in audio description (AD) guidelines. However, some scholars have pointed to the advantages of a more subjective, creative, or narrative rendition. In this paper, we seek to determine to what extent subjectivity and objectivity coexist in Spanish filmic AD. In order to fulfil our aim, we operationalised objectivity by means of multimodal analysis based on Chaume's (2004) classification of meaning codes. To operationalise objectivity through multimodality, we employed a qualitative content analysis and examined whether visual and acoustic information was rendered objectively or subjectively in a corpus of four Spanish AD scripts from Netflix. Our results show, firstly, that objectivity and subjectivity interact in the mobility, iconographic, and editing codes. Moreover, dissimilarities arise in the way these meaning codes are described in Spanish: movement is mainly described objectively, whereas iconography and editing are rendered either subjectively or objectively. In conclusion, we can state that neither is objectivity systematically applied in our AD scripts, nor is a purely narrative AD to be found, but rather a mixture of both coexists. Furthermore, this coexistence seems unequal, since different tendencies can be identified in the way that movement, editing, and iconography are described.

**Key words:** audiovisual translation, accessibility, audio description, operationalisation, objectivity, multimodality, meaning codes.

## Introduction

Audio describers do not usually create audio description (AD) scripts by merely following their intuition of what is appropriate. In fact, many countries and video on demand (VOD) platforms have guidelines which provide recommendations about what to describe and how to do so. Among the many proposals that can be found in these documents, objective description is recurrently encouraged. Despite the ubiquity of objectivity, some audiovisual translation (AVT) and media accessibility (MA) scholars have explored the feasibility and advantages of creating alternative versions of conventional objective AD, namely through more creative, subjective, and narrative descriptions.

Resulting from this ongoing debate, this paper moves beyond the possibility of creating a purely narrative or subjective AD and attempts to answer a different research question: do subjectivity and objectivity already coexist in AD scripts? And if indeed they do, then current filmic AD would be the result of combining subjective and objective elements. With the aim of determining this interaction between subjectivity and objectivity, we resorted to the hypothesis that multimodal analysis based on meaning codes will be an unexplored yet appropriate way of operationalising objectivity (and therefore also subjectivity) to ascertain how the two manifest in AD scripts.

Throughout the paper, we first establish the relationship between multimodality and AD by providing a brief overview of multimodality as a discipline as well as its association with AVT, particularly with AD. After that, some space is devoted to revising key proposals related to the objectivity versus subjectivity debate in AD. Then a full explanation of the descriptive methodology is provided, specifically focusing on the corpus compilation stages, the multimodal application of meaning codes by means of qualitative content analysis (QCA), and the subsequent operationalisation of objectivity. After that, the results derived from the analysis are thoroughly detailed and followed by the discussion section and a series of concluding remarks.

### 1. Multimodality and AD

When defining multimodality, the concept of “semiotic modes” and their interrelation are usually brought up. For instance, Kress and Van Leeuwen (2001) define multimodality as the use of semiotic modes and the way these modes combine in a product or event. Similarly, Stöckl (2004, p. 9) highlights that “multimodal refers to communicative artefacts and processes which combine various sign systems (modes) and whose production and reception calls upon the communicators to semantically and formally interrelate all sign repertoires present.”

More precisely, semiotic modes could be described as “semiotic resources”, whose meaning is “culturally made, socially agreed and socially and culturally specific” (Kress, 2014, p. 60). As for the interrelation of modes, it is precisely because of this interaction that meaning is created (O’Halloran et al., 2015). Consequently, multimodality focuses on how semiotic modes combine to constitute the

multimodal text (Taylor, 2020). In this vein, Villanueva-Jordán (2021) points out four key notions involved in the generation of meaning in multimodal texts: the resource integration principle, multiplying meaning, modal density, and modal configuration.

On the one hand, the resource integration principle alludes to the way modes relate to each other (Baldry & Thibault, 2006). On the other hand, Lemke states that the meanings from one set of semiotic modes are affected by the meanings resulting from other semiotic modes, “thus multiplying the set of possible meanings that can be made” (1998, p. 92). Finally, Norris (2004) refers to “modal density” as the intensity and complexity achieved by means of the interaction of different modes and, due to this density, some modes might become more relevant than others, a hierarchy called “modal configuration” (Norris, 2004).

Having defined the basic traits of multimodality, it is worth examining the main lines of its development as a discipline. An early period of multimodality can be traced back to the late 20<sup>th</sup> century (Taylor, 2020), when multimodal research focused on visual images, particularly in the works of Kress and Van Leeuwen (1996), Baldry (2000), and Thibault and Baldry (2001). However, this tendency changed during the 21<sup>st</sup> century (Taylor, 2020), when multimodality moved from an approach based on Halliday’s theory of language as a social semiotic (O’Halloran, 2012) to an exploration of proposals from other disciplines, such as film studies, a combination further developed by AVT scholars.

It is important to note that the relationship between multimodality and AVT depends on a third discipline: semiotics. Kourdis (2022) affirms that semiotics constitutes the basis of multimodality and multimodality is in turn one of the theoretical frameworks used by AVT scholars. As a matter of fact, when it comes to defining AVT, particularly the audiovisual text, it is usually done in semiotic and multimodal terms. For instance, Díaz-Cintas (2020) states that audiovisual texts are semiotic composites in which codes merge to create meaning. Among the scholars who have studied the characteristics of the audiovisual text, Zabalbeascoa (2008) proposed a classification of two parameters specifically affecting AVT: channels of communication (audio and visual) and sign codes (verbal and non-verbal).

Another AVT scholar who has delved into the semiotic complexity and multimodal essence of the audiovisual text is Chaume (2004, 2012), who goes beyond the Hallidayan model and proposes an analysis based on eleven filmic meaning codes transmitted either by the acoustic or by the visual channel that specifically affect AVT (Table 1). These codes are constituted by signs and their meaning is conventionalised by culture, which relates to the traditional multimodal notion of the “semiotic mode” (Villanueva-Jordán, 2021). As for the acoustic codes, the linguistic code refers to any language information appearing in dialogues, monologues, a narrator’s voice, etc. The paralinguistic code has to do with aural nonverbal information. 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 (e.g., a doorbell ringing). Finally, the sound position code deals with the origin of characters’ voices. Moving on to the visual codes, the iconographic code is composed of

indexes, icons, and symbols. Conversely, the photographic code alludes to information related to perspective, light, and colour. 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 (e.g., a close-up shot to focus attention on a character's face). The graphic code deals with written language appearing on screen, such as titles, subtitles, credits, etc. Finally, the editing code has to do with the film transition marks that organise the film in shots, sequences and so forth.

**Table 1**

*Chaume's Classification of the Meaning Codes*

<b>Acoustic Channel</b>	<b>Visual Channel</b>
Linguistic code	Iconographic code
Paralinguistic code	Photographic code
Musical code	Mobility code
Special effects code	Shot code
Sound position code	Graphic code
	Editing code

*Source:* Chaume, 2012, p. 172.

Having related multimodality and AVT, we should look at how this relationship applies to AD. AD could be defined as an AVT mode that aims to make contents accessible primarily for the blind and visually impaired. Vercauteren states that AD translates “the visual and aural elements that [people with sight loss – emphasis added] do not have access to into a verbal commentary” (2022, p. 78).

As previously indicated, there is a link among semiotics, multimodality, and AVT (and thus AD). Of the three disciplines, AD and semiotics have long been related by scholars. To better understand this connection, we must first turn to Jakobson (1959), who made a distinction between intralingual translation, interlingual translation, and intersemiotic translation (the translation of verbal signs into a nonverbal sign system). Consequently, many scholars have linked the concept of intersemiotic translation to AD (Matamala, 2019; Taylor, 2020), although following the inverse order: AD translates from a nonverbal sign system into a verbal sign system.

Apart from the semiotic nature of AD, a connection between multimodality and AD can also be made: “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 authors have already explored the association between multimodality and AD, such as Braun (2011), Reviers (2018), and Matamala (2019). 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). Reviere (2018) employs a corpus-based multimodal study to describe linguistic features of AD scripts and the role they play in the communicative function. Finally, Matamala (2019) developed the VIW (Visual Into Words) project, a multimodal and multilingual AD corpus comprising the AD in English, Spanish, and Catalan of a single short film where the linguistic and film levels were studied by means of the corpus analysis tool ELAN.

## 2. Objectivity in AD and Alternatives

If we were to focus on the notion of objectivity in AD, opposite perspectives collide. On the one hand, AD guidelines used by many countries and VOD platforms prefer the traditional perspective of maintaining an objective or denotational description. On the other hand, growing interest has developed among researchers towards an alternative to objectivity, with some even advocating for the advantages of a more narrative, subjective, or interpretative AD.

Regarding the traditional approach, this is the stance supported by guidelines followed in countries like Spain, France, and the UK. Spain's *Norma UNE 153020* states: "Do not describe what is easily extracted from the work nor any subjective points of view" (AENOR, 2005, p. 8). Similarly, France's *La Charte de l'audiodescription* suggests: "The description must be done in an objective way so as not to impose your own feelings, but to provoke them" and "do not use subjective adjectives except when the characteristic is evident" (Morisset & Gonant, 2008, pp. 2–4). In the UK, the 2015 version of Ofcom still recommended: "The description should only provide information about what can be seen on the screen" and "adverbs are a useful shorthand to describing emotions and actions but should not be subjective" (2015, p. 21). We could also include Joel Snyder's *The Visual Made Verbal*, where he asserts: "The best audio describer is sometimes referred to as a *verbal camera lens*, objectively recounting visual aspects of an event"; "it's critical to maintain that sense of objectivity"; "the oft-referenced first rule of description is to *Describe what you see*" (2014, pp. 163–164). Finally, another case of objectivity support is Netflix, whose guidelines state: "Description should be factual" and "description should not be opinionated" (2022, pp. 1–2).

On the other hand, we have another approach supported by some AVT and MA scholars, who acknowledge that the ubiquity of objectivity in AD is problematic, and they even explore the advantages of a narrative, subjective, or interpretative AD as an alternative. Accordingly, Jankowska states: "A hundred per cent objectivity is impossible, and AD is always subjective, since it is a choice made by a particular audio describer" (2015, p. 22), just as Holsanova supports the idea that: "An 'objective' AD is problematic considering the range of individual and contextual factors that influence how we focus our attention, select relevant information, and describe a scene" (2016, p. 53). Beyond the rejection of total objectivity in AD, Kruger (2010) was the first to introduce the notion of the "descriptive-narrative continuum", where he proposes three types of AD: an explicitly descriptive AD that focuses on the visual aspects to be described that would be objective, an AD with some narrative markers and subjective interpretation located somewhere in the middle of the continuum, and finally

an audio narration that favours a coherent narration in detriment of what can be seen on the screen. Another important contribution is Szarkowska's (2013) proposal of *auteur description*, which includes the director's creative vision in the AD by means of vivid and emotional language, additional information about the characters, their emotions, as well as actions and settings not necessarily visible on screen. Similarly, Jankowska (2015) alludes to a creative audio description that includes colour 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. Fryer and Freeman (2012) propose a style called "cinematic AD" that includes filmic language. Finally, Bardini (2020) distinguishes three AD styles: conventional, cinematographic, and narrative. Conventional AD is a denotational and objective description of what is on screen, cinematographic AD is interpretative and uses and interprets filmic terms, whereas narrative AD interprets information through vivid verbal elements that create a fluid and coherent narration.

Consequently, there seem to be two different trends regarding the contents of AD. On the one hand, there is a conventional trend, usually suggested by guidelines, which encourages the use of objectivity to denotatively describe what is on screen. On the other hand, an alternative scholarly trend rejects the idea of complete objectivity being possible and explores the advantages of a narrative, cinematic, interpretative, or creative AD.

### 3. Methodology

As mentioned before, this paper focuses on whether subjectivity and objectivity coexist in AD scripts. To ascertain this coexistence, the hypothesis followed is that multimodal analysis based on meaning codes will be an appropriate way of operationalising objectivity and subjectivity. The reason behind the use of multimodality is that, although some authors have already explored this path, we believe that some new steps can be taken into inspecting meaning codes because they are a well-known classification in AVT (which could promote research combined with other AVT modes) and because meaning codes allow an in-depth analysis of very specific pieces of information.

Therefore, we resorted to a descriptive methodology that drew on a corpus compiled as a result of three levels of selection criteria. The first level led to corpus 0, which consisted of a classification of the most appropriate audiovisual material. The second level used corpus 0 to create corpus 1, where the most convenient series from corpus 0 were selected. Finally, the most representative fragments from corpus 1 were chosen in the third level, which resulted in corpus 2. After that, the multimodal analysis based on meaning codes was applied to corpus 2 following the principles of QCA so that the interaction between objectivity and subjectivity could be clearly evaluated.

### 3.1. Corpus 0

To properly portray the usual multimodal contents of AD scripts to then be analysed in terms of objectivity-subjectivity, a descriptive study based on a corpus was considered the best methodological option. As the basis of the analysis, the corpus selection required rigorous principles that would enable possible future replication. 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 provided us with “corpus 0”, a database consisting of 23 Netflix series from 2021 with Spanish AD.

#### 3.1.1. Availability

An AD corpus-based study requires long-term access to a considerable number of audio described resources. For this reason, it was essential to opt for a source that would provide us with enough material for as long as possible. Consequently, VOD platforms were thought to meet that requirement. The main options were four well-known sites: Netflix, Amazon Prime, Disney+, and HBO Max. Despite being renowned VOD platforms, they featured a dissimilar quantity of accessible content. On the one hand, HBO Max had no audio described products in Spain from 2021 to the best of our knowledge. On the other hand, Disney+ and Amazon Prime had no way of searching for specific audio described material. Ultimately, we chose Netflix because the catalogue options allow users to retrieve a reasonable number of audio described films and series.

#### 3.1.2. Production

With Netflix as the audiovisual source, it was necessary to decide which type of productions available on the platform would be most appropriate: films, series, documentaries, etc. Among the array of possibilities, we opted for series as the best option due to their genre diversity, their vast presence on the platform, their popularity, and the frequency with which they had AD. Certainly, most of these factors could also be attributed to films, but another reason behind the choice was the amount of time with audio described content. While an average film might provide approximately 120 minutes of material containing AD, the third season of the series *Money Heist* consists of 10 episodes, each of them being around 45 minutes long, which provides 450 minutes of content containing AD.

#### 3.1.3. Linguistic

Following the purpose of studying the AD scripts in Spanish, we selected those series which included Spanish AD. The resulting options mainly consisted of series from Spanish-speaking countries, such as Spain or Mexico. Moreover, the catalogue also provided us with series from other countries (like France, South Korea, or Italy) that had Spanish AD along with that in other languages.

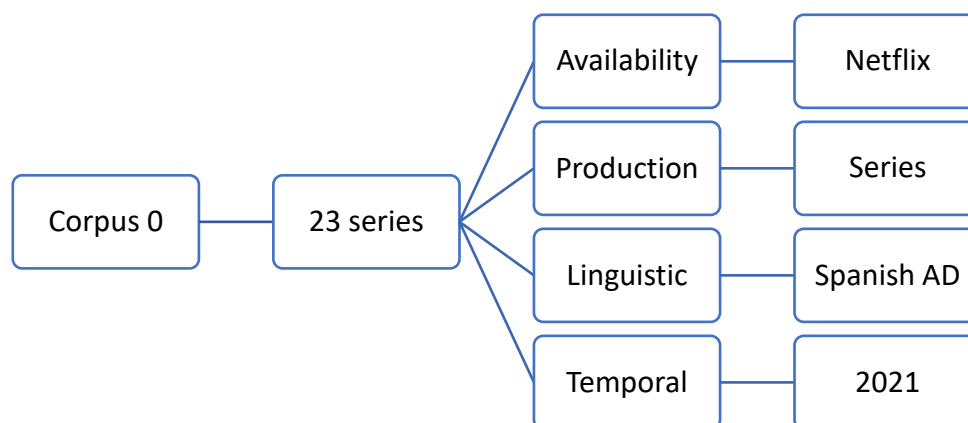
### 3.1.4. Temporal

The catalogues of VOD platforms are dynamic, with parts of their content constantly being added or removed. Therefore, a temporal criterion was established to restrict the selection to the most recent series. Since this research was conducted in 2021, we focused on series with at least one season released that year.

As shown in Figure 1, after applying these four filters, we obtained corpus 0 comprising 23 Netflix series from 2021 with Spanish AD.

**Figure 1**

*Corpus 0 Criteria*



*Source:* Author's own elaboration

### 3.2. Corpus 1

Analysing all episodes of 23 series seemed to be an excessive task, so drawing on our corpus 0, we configured a further refined database called "corpus 1", whose aim was to focus on an appropriate number of series with common features. To compile corpus 1, we established another three selection criteria: series origin, date, and genre. Consequently, the application of these criteria led us to corpus 1 comprising four Spanish drama and thriller series with seasons released on Netflix in 2021.



### 3.2.1. Origin

The audiovisual material from corpus 0 included series from many countries, so the possibilities were limited to one place of origin for the sake of accuracy. Thus, we considered the most recurrent countries in corpus 0: Spain, the UK, and the USA. The number of American series that we had access to from Spain was considerable, yet at the same time lacking homogeneity, with contents ranging from animation to documentaries. On the other hand, the UK series were more uniform in their content, but they were fewer in number. However, in corpus 0, Spanish series were frequent, and their characteristics were similar, so Spain was selected as the most suitable origin for corpus 1.

### 3.2.2. Date

Although the previous temporal cut-off privileged series from 2021, some of our options from corpus 0 also had previous seasons, such as *Money Heist*, with 48 episodes over three seasons spanning from 2017 to 2021. Accordingly, corpus 1 was restricted to those Netflix seasons which were released in 2021 to portray the most updated characteristics of AD scripts.

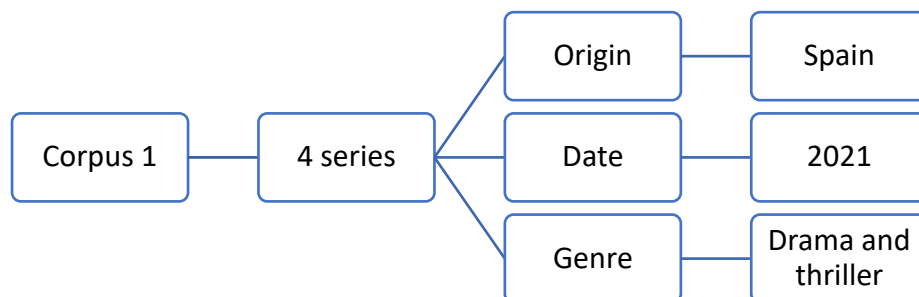
### 3.2.3. Genre

Seeking to obtain robust results, it seemed more coherent to concentrate on a single genre out of the many possibilities we had in corpus 0. This way, we ensured the resulting data would be homogenous, which at the same time enabled any hypothetical tendency to be more rigorous. However, assigning a single genre to a particular series can be reductive, and sites like IMDb or Netflix itself normally group series or films under more than one genre. Regarding corpus 0, drama was the most frequent genre, but it was also part of a number of other genres like musical, adventure, or biographical series that was too broad. Consequently, the search was further restricted to those series which belonged to both drama and the second most recurrent genre in the corpus, i.e., thriller.

The application of these three criteria to corpus 0 resulted in corpus 1 (Figure 2), containing four Netflix Spanish drama and thriller series with seasons released in 2021: *Elite*, *Elite: Short Stories*, *Money Heist*, and *Sky Rojo*.

**Figure 2**

*Corpus 1 criteria*



*Source:* Author's own elaboration.

### **3.3. Corpus 2**

Even after having created the more limited corpus 1, there was still too much material, so it was fundamental to establish corpus 2 with some sort of filter to select which episodes and scenes to consider for the subsequent multimodal analysis. Ultimately, we decided to use mainly the first episodes from the latest seasons of our four series from corpus 1, and then we limited the analysis to approximately the first five minutes of those episodes.

#### **3.3.1. Episode Selection**

With the aim of avoiding any arbitrary corpus choice, we adopted the simple rule of considering the first episode from each latest season as the ideal option. When previous seasons exist, first episodes usually allude to important past events that the audience should keep in mind, which means that there will be important information that needs to be audio described. On the other hand, when series start from scratch, the first episode is particularly interesting because it establishes the essential information: the main characters, places, plot elements, etc. It must be noted that this principle could not be applied in some of the series due to scene selection restrictions, as will be explained in the next section.

#### **3.3.2. Scene Selection**

Having the first episode as the ideal option, we implemented another criterion to help select the most convenient fragments. Thus, we concur with Remael and Vercauteren (2007), who stated that the first ten minutes of a film determine the audience's expectations of what the remainder of the

film will be. In addition, during that period of time, we can expect to be introduced to some key elements that need to be audio described, like characters, places, or the main action. However, we only partially employed Remael and Vercauteren’s criterion. Unlike in films, ten minutes constitute a considerable excerpt of an episode, so only five minutes were considered. In other words, we selected approximately the first five minutes of the first episode from the latest season of each of our series, starting from the very beginning up to an appropriate scene or sequence change. However, scene selection considerably restricted our previous first-episode principle. Dialogues were too dense and with virtually no room for AD in the first episodes of *Elite: Short Stories* and *Money Heist*, so in those cases we maintained the first-five-minute principle but chose another episode (Table 2).

As Figure 3 illustrates, we used corpus 1 to select, when possible, the first episode from the latest season of our four series. Once this process was complete, we limited the multimodal analysis to approximately the first five minutes of those episodes.

**Table 2**

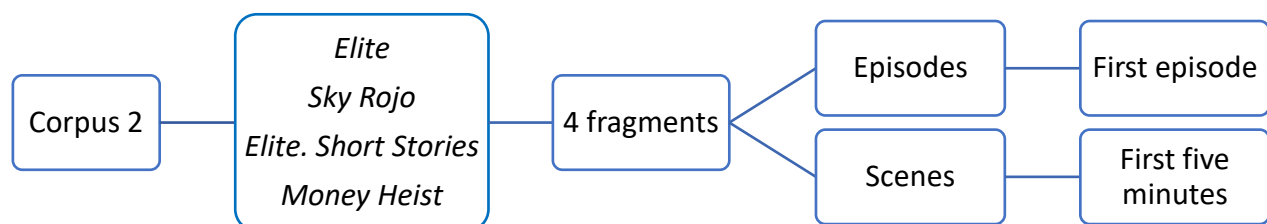
*Episode and Scene Selection*

Serie	Episode	Scene
Elite	E1S4 The New Order	00:00–04:00
Sky Rojo	E1S2 Hookers Didn’t Kiss on the Lips	00:00–06:34
Elite. Short Stories	E3S1 Guzmán Caye Rebe. Part 3.	00:00–04:47
Money Heist	E2S3 Do You Believe in Reincarnation?	00:00–05:54

Source: Author’s own elaboration.

**Figure 3**

*Corpus 2 Criteria*

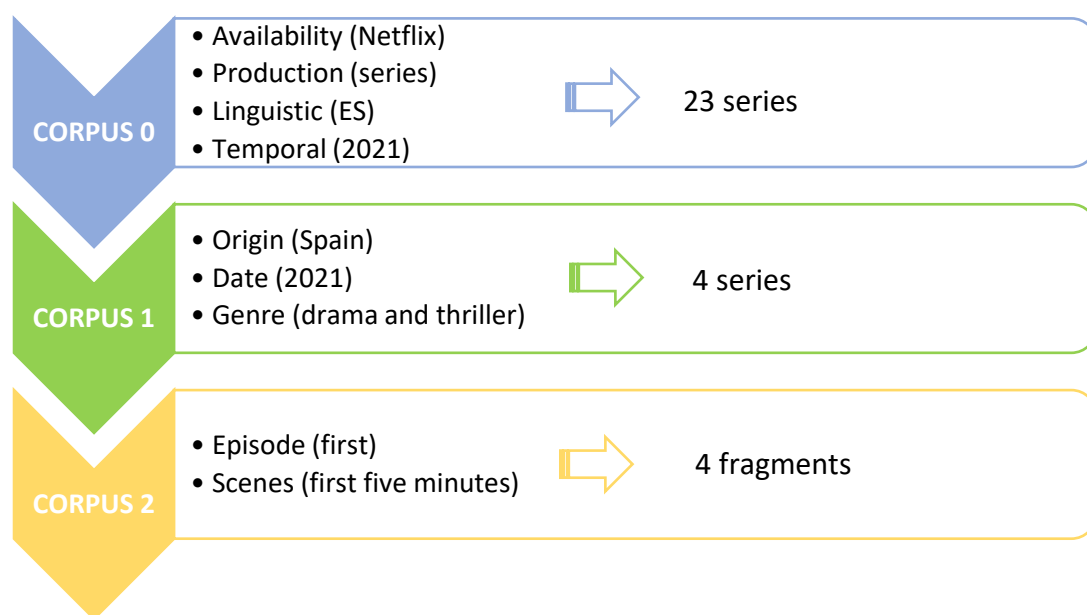


Source: Author’s own elaboration.

To summarise the corpus compilation procedure, Figure 4 gives an overview of its three main stages. First, we created corpus 0 consisting of a classification of what audiovisual products would be appropriate. Stemming from corpus 0, corpus 1 was designed with those series considered to be the most convenient. Finally, corpus 1 gave rise to corpus 2 with the most suitable fragments to be tackled in the multimodal analysis.

**Figure 4**

*Corpus Compilation Stages*



*Source:* Author's own elaboration.

### 3.4. Multimodal Analysis

Once the four fragments were selected, we transcribed them. In order to do so, we did not limit our transcriptions to the AD excerpts, but we also included dialogues so that the interaction between AD and other semiotic information could be easily examined. It must be acknowledged that the resulting layout (Table 3) is influenced by other researchers' transcription patterns, like Walczak's (2017). For communicative purposes, a translation of Table 3 is provided in Table 4.

**Table 3**

*AD Transcription Layout*

---

**Élite (E1S4 – «El nuevo orden»)**

---

[54:33] Una letra N roja se despliega y adopta varios colores.  
[54:29] «Una serie original de Netflix». «Una producción de Zeta Studios». En el Club del Lago.  
**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 se para delante de Guzmán. El cuerpo de una chica morena de pelo corto con un vestido blanco flota boca arriba en el lago.  
[53:47] Sobre el lago los fuegos artificiales iluminan el cielo nocturno.  
**POLICÍA:** Guzmán.

---

*Source:* Author's own elaboration

**Table 4**

*AD Transcription Translated into English*

---

**Elite (E1S4 – “The New Order”)**

---

[54:33] A red letter N unfolds and takes on several colours.  
[54:29] “A Netflix original series”. “A Zeta Studios production”. In the Lake Club.  
**POLICE OFFICER:** Guzmán. Guzmán. I need to ask you a few questions. What happened? I need you to tell me what you have seen tonight. What happened?  
[54:02] A police officer stands before Guzmán. The body of a short dark-haired girl in a white dress floats face up in the lake.  
[53:47] Fireworks light up the night sky over the lake.  
**POLICE OFFICER:** Guzmán.

---

*Source:* Author's own elaboration.

Having transcribed the dialogues and AD fragments, we then applied the multimodal analysis based on Chaume's (2004) meaning codes. As stated before, some of these codes are transmitted through the visual channel and others through the acoustic channel. Consequently, as AD is primarily aimed at compensating the loss of visual information, we might assume that acoustic codes would be semiotically redundant. Nevertheless, all the codes were incorporated in the analysis to fully characterise any possible multimodal component of the AD script.

In order to associate the AD information with the meaning codes, we applied the principles of 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 et al., 2011, p. 138). Thus, Chaume’s classification of the meaning codes was adapted to the characteristics of AD and the resulting combination constituted the codebook for the analysis (Table 5).

**Table 5**

*Codebook of Meaning Codes Applied to AD*

Channel	Code	Symbol	Definition	AD example <sup>1</sup>
ACOUSTIC CHANNEL	linguistic	(L)	Presence of any information related to language: morphology, syntax, typology, language variation, etc.	– Subtitles appear as a reporter speaks <u>in Japanese</u> .
	paralinguistic	(P)	Inclusion of aural non-verbal information, such as voice rhythm, prosody, volume, tone, etc.	– The protesters <u>cheer</u> .
	musical	(M)	Description of signs that can be attributed to music, such as singing, music genres, songs, etc.	– He <u>sings flamenco</u> .
	special effects	(SE)	Presence of signs related to sound elements not uttered by characters.	– A <u>fireball</u> consumes the van.
	sound position	(SP)	Inclusion of any information pointing to the origin of a particular sound in the AD.	– The Professor grabs his <u>radio mouthpiece</u> .
VISUAL CHANNEL	iconographic	(I)	Reference to indexes, icons, and symbols.	– <u>A red letter N</u> unfolds into a spectrum of colours.
	photographic	(PH)	Allusion to perspective, light, or colour in the AD.	– <u>Light shines</u> on the edges of clouds.
	shot	(SHT)	Description of any information related to camera movements and angles, as well as the information they convey.	– <u>In slow motion</u> , Raquel leads Mónica towards the front door.
	mobility	(MB)	Presence of proxemic, kinesic, or mouth articulation signs in the AD.	– Rebeka <u>chases</u> Guzmán.
	graphic	(G)	Inclusion of titles, intertitles, texts, or subtitles.	– <u>A Netflix series</u> .
	editing	(ED)	Information about audiovisual punctuation signs, such as temporal or physical scene progressions.	– <u>Now</u> , a girl in a flowing blue dress floats face-up <u>in the lake</u> .

Source: Author’s own elaboration.

<sup>1</sup> In order to better illustrate how meaning codes appear in AD scripts, original English AD fragments from the selected series have been used here, since Netflix has AD in both languages for these series. However, the analysis has been carried out using the original Spanish AD, as Table 6 shows.

To easily identify the meaning codes in the AD script, each code was allotted a symbol consisting of its first letter(s). In this way, (I) would stand for a sign from the iconographic code. The whole AD transcription corpus was analysed by cataloguing every sign (usually at the clause level) as instances belonging to any of the eleven meaning codes. The next step was deciding whether signs were described in an objective or subjective way in order to operationalise objectivity.

With the aim of establishing a coherent differentiation, we created a series of objectivity and subjectivity parameters drawing on the guidelines and researchers' proposals from section 2. Thus, a sign was considered to be objectively rendered when there was no subjective interpretation or point of view (AENOR, 2005; Morisset & Gonant, 2008; Netflix, 2021), no references to the describer's feelings (Morisset & Gonant, 2008), when it just included what could be seen on screen (Ofcom, 2021; Snyder, 2014; Bardini, 2020), and when it used no subjective adjectives (Ofcom, 2021), just factual descriptions (Netflix, 2021). Conversely, a sign was regarded to be subjectively rendered when there was any sort of interpretation (Kruger, 2010; Bardini, 2020), if it contained emotional language (Szarkowska, 2013), additional information about actions or settings (Szarkowska, 2013), vivid language (Jankowska, 2015; Walczak, 2017; Bardini, 2020), it resorted to metaphors (Jankowska, 2015), it contained references to emotions (Jankowska, 2015; Walczak, 2017), or it included film terminology (Jankowska, 2015; Walczak, 2017; Fryer & Freeman, 2012; Bardini, 2020).

To make the difference clearer, every code symbol was followed by an "O" in red or "S" in blue, which stand for "objective" or "subjective", respectively. In this way, (I\_O) would stand for an objective sign from the iconographic code, whereas (I\_S) would refer to a subjective counterpart. An example of the AD multimodal analysis can be observed in Table 6 (translated in Table 7), whereas Tables 8 and 9 depict a more developed version of the codebook in terms of objectivity and subjectivity.

**Table 6**

*Excerpt From the AD Multimodal Analysis*

---

**Elite (E1S4 – “The New Order”)**

---

[54:33] A red letter N (I\_O) unfolds (MB\_O) and takes on several colours (PH\_O).

[54:29] “A Netflix original series” (G\_O). “A Zeta Studios production” (G\_O). In the Lake Club (ED\_S).

**POLICE OFFICER:** Guzmán. Guzmán. I need to ask you a few questions. What happened? I need you to tell me what you have seen tonight. What happened?

[54:02] A police officer (I\_S) stands before Guzmán (MB\_O). The body of a short dark-haired girl in a white dress (I\_S) floats face up (MB\_O) in the lake (ED\_O).

[53:47] Fireworks light up (I\_O) the night sky (ED\_O) over the lake (ED\_O).

**POLICE OFFICER:** Guzmán.

---

*Source:* Author’s own elaboration.

**Table 7**

*Excerpt From the AD Multimodal Analysis Translated Into English*

---

**Élite (E1S4 – «El nuevo orden»)**

---

[54:33] Una letra N roja (I\_O) se despliega (MB\_O) y adopta varios colores (PH\_O).

[54:29] «Una serie original de Netflix» (G\_O). «Una producción de Zeta Studios» (G\_O). En el Club del Lago (ED\_S).

**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\_S) se para delante de Guzmán (MB\_O). El cuerpo de una chica morena de pelo corto con un vestido blanco (I\_S) flota boca arriba (MB\_O) en el lago (ED\_O).

[53:47] Sobre el lago (ED\_O) los fuegos artificiales iluminan (I\_O) el cielo nocturno (ED\_O).

**POLICÍA:** Guzmán.

---

*Source:* Author’s own elaboration



**Table 8**

*Codebook of Objective and Subjective Acoustic Codes*

Channel	Code	Description	Symbol	Definition
ACOUSTIC	Linguistic	Objective	(L_O)	Denotational information related to language.
		Subjective	(L_S)	Interpretation of linguistic information, such as origin of an accent or what language someone speaks.
	Paralinguistic	Objective	(P_O)	Inclusion of aural non-verbal information.
		Subjective	(P_S)	Use of filmic terms (“canned laughter”), vivid language, similes (“laugh like a hyena”), or metaphors to describe non-verbal information.
	Musical	Objective	(M_O)	Basic elements attributed to music in cases where this information is essential.
		Subjective	(M_S)	Additional information about music genre, associated emotions (“sad music”), interpretation of the author or song title, or use of filmic terms (“diegetic music”).
	Special effects	Objective	(SE_O)	Presence of signs related to sound elements not uttered by characters when this information is difficult to identify.
		Subjective	(SE_S)	Cinematographic terms (“stop motion”) or use of metaphors or similes (“firecrackers like a bomb”).
	Sound position	Objective	(SP_O)	Inclusion of any information pointing to the origin of a particular sound in the AD when it is essential.
		Subjective	(SP_S)	Additional information, interpretation of any effect this sound provokes, or use of vivid language to describe it.

Source: Author’s own elaboration.

**Table 9**

*Codebook of Objective and Subjective Visual Codes*

Channel	Code	Description	Symbol	Definition
VISUAL	Iconographic	Objective	(I_O)	Reference to indexes, icons, and symbols without explaining their meaning.
		Subjective	(I_S)	Explanation of index, icon, and symbol meaning. A vivid description or use of metaphors and similes.
	Photographic	Objective	(PH_O)	Denotational allusion to perspective, light, or colour.
		Subjective	(PH_S)	Use of filmic language (“bounce lighting”), vivid or metaphorical language, or interpretation of photographic meaning.
	Shot	Objective	(SHT_O)	Denotational description of camera movements and angles.
		Subjective	(SHT_S)	Interpretation of shot meaning, use of metaphors, extensive information about it, filmic terms about shot typology, etc.
	Mobility	Objective	(MB_O)	Proxemic, kinesic, or mouth articulation signs with no interpretation of their meaning.

Channel	Code	Description	Symbol	Definition
Graphic		Subjective	(MB_S)	Interpretation of movements, vivid descriptions, or use of rhetorical figures.
		Objective	(G_O)	Information about title, intertitle, text, or subtitle content as part of the AD unit (“A Netflix series”).
		Subjective	(G_S)	Emphasis on the fact that a text is being read (“Text appears”) or interpretation of its contents (“a text in French”).
		Objective	(ED_O)	Information about audiovisual punctuation signs, such as temporal or physical scene progressions.
Editing		Subjective	(ED_S)	Filmic expressions (“fade to black”), vivid language about space and time changes, interpretation of these changes, or use of metaphorical language.

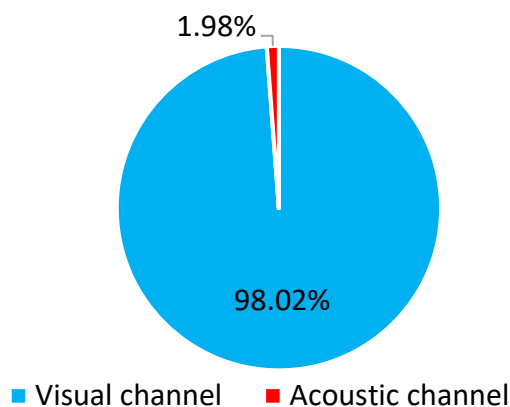
Source: Author’s own elaboration.

#### 4. Results

Although QCA is a qualitative approach, it was turned into a mixed-method approach by quantitatively measuring which channel constituted most of the AD script or how frequently signs from each code appeared, i.e., the multimodal configuration. Regarding how well-proportioned AD scripts are in terms of channels, the quantitative results obtained by our qualitative analysis show a sharp contrast between the acoustic and visual codes (Figure 5). The selected AD scripts contained 252 signs, of which 247 belonged to the visual channel (98.02%), leaving the remaining 5 signs to the acoustic channel (1.98%).

Figure 5

Channel Proportion in the AD Script



Source: Authors’ own elaboration.

A closer examination of the data (Table 10) allows us to determine the exact proportion of every code. As for the acoustic channel, there have only been 3 signs found from the linguistic code (1.2% of the AD script), 1 sign from the paralinguistic code (0.39%) and 1 sign from the special effects code (0.39%). As for the visual codes, we found 111 signs from the mobility code (44.05%), 75 signs from the iconographic code (29.76%), 31 signs from the graphic code (12.3%), 23 signs from the editing code (9.13%), and 7 signs from the photographic code (2.78%).

**Table 10**

*Meaning Codes Results*

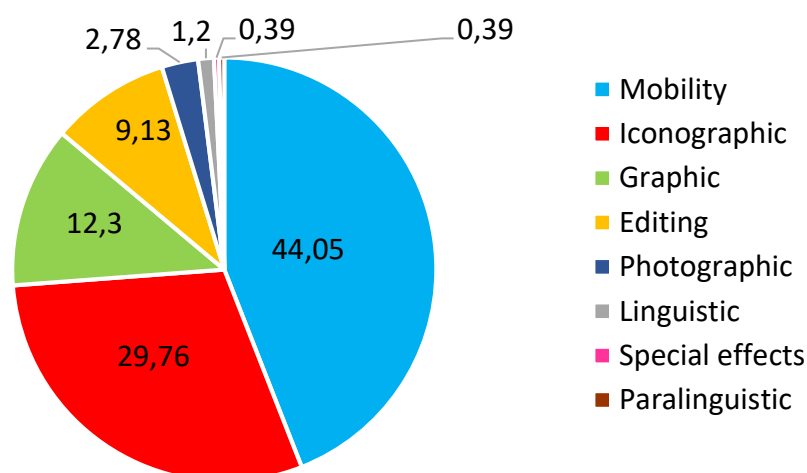
Acoustic channel	Visual channel
Linguistic (L): 1.2%	Iconographic (I): 29.76%
Paralinguistic (P): 0.39%	Photographic (PH): 2.78%
Musical (M): 0%	Mobility (MB): 44.05%
Special effects (SE): 0.39%	Shot (SHT): 0%
Sound position (SP): 0%	Graphic (G): 12.3%
	Editing (ED): 9.13%

Source: Author's own elaboration.

Consequently, the quantitative data obtained by means of QCA allows us to describe the multimodal configuration of the AD script as represented in Figure 6.

**Figure 6**

*Multimodal Configuration of the AD Script*

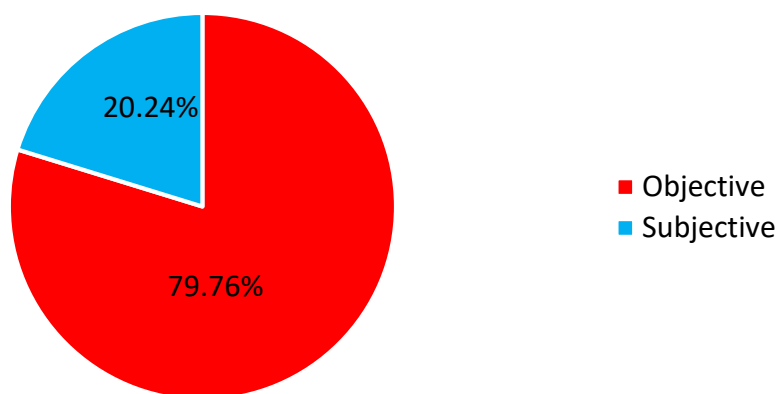


Source: Author's own elaboration.

Moreover, multimodality allows us to operationalise objectivity by ascertaining which codes tend to be rendered objectively or subjectively, how frequently this occurs, whether subjectivity and objectivity coexist in certain codes, etc. Thus, Figure 7 shows the proportion of objective (79.76%) and subjective (20.24%) signs throughout the AD scripts. Then, Table 11 illustrates how many signs from the four main meaning codes that comprise the multimodal configuration of AD are rendered in an objective or subjective way and what percentage of the whole code this rendition represents. In the same vein, Figure 8 depicts the interaction between objectivity and subjectivity in the multimodal configuration of AD.

**Figure 7**

*Proportion of Objective and Subjective Information in AD Scripts*



Source: Author's own elaboration

**Table 11**

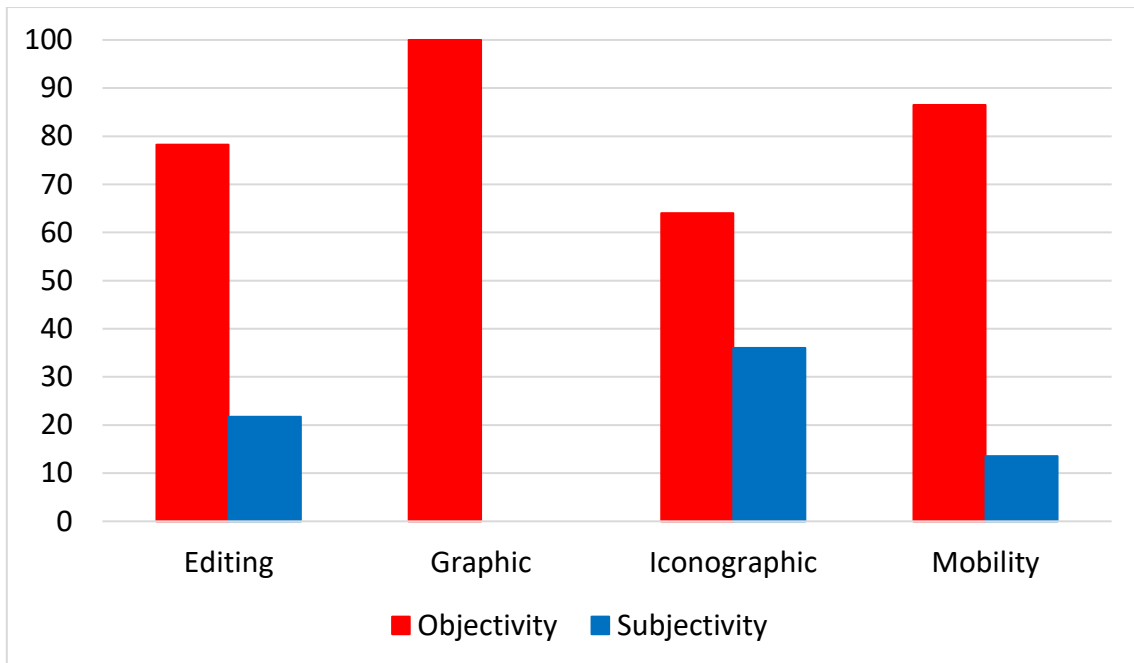
*Objectivity and Subjectivity in AD*

Code	Description	Signs	Percentage
Mobility (MB)	Objective	96	86.48%
	Subjective	15	13.52%
Iconographic (I)	Objective	48	64%
	Subjective	27	36%
Graphic (G)	Objective	31	100%
	Subjective	0	0%
Editing (ED)	Objective	18	78.26%
	Subjective	5	21.74%

Source: Author's own elaboration.

**Figure 8**

*Coexistence of Objectivity and Subjectivity in AD*



*Source:* Author's own elaboration.

## 5. Discussion and Conclusions

Drawing on the results obtained through **the** multimodal analysis, we can state that meaning codes have enabled the operationalisation of objectivity in these scripts.

Unsurprisingly, one of the most remarkable tendencies is the absolute lack of information balance in terms of channels (Figure 5). Given the scarce presence of acoustic codes in the analysed fragments, we can safely say that our AD scripts are essentially composed of visual information. Far from undermining the analytical relevance of multimodality, this semiotic disproportion has a multimodal explanation. Considering that AD primarily intends to compensate for any restricted access to visual elements, there should not be any need to also compensate for aural information unless it is unclear from context, since most primary AD users would have access to the acoustic channel. Therefore, acoustic elements might be considered semiotically redundant. However, our data points at a minor presence of acoustic codes. Rare as they might be, they still comprise a fraction of the AD, which supports the need for the fully articulated multimodal approach provided by all meaning codes.

Conversely, another tendency is that virtually the entire configuration of the AD script is constituted by visual elements (Figure 6). This shares the same multimodal explanation mentioned previously. Within the visual channel, further regularities are observed. Mobility and iconographic codes encompass almost three quarters of the whole script. Following these, graphic and editing codes

form a considerable section of the AD script, but to a lesser extent. Finally, other codes such as the photographic code are uncommon.

All in all, it can be stated that the core of AD multimodal configuration revolves around just four codes (mobility, iconographic, editing, and graphic), which account for nearly 95% of the AD script (Table 10 and Figure 6). This leads to the main research question: do objectivity and subjectivity coexist in these codes? If we analyse the AD scripts, around 80% of the information tends to be rendered objectively, whereas approximately 20% of the scripts resort to subjective elements (Figure 7). Furthermore, the multimodal configuration reveals that three of the four main codes present subjective elements in different proportions, whereas the graphic code tends to only be rendered objectively (Table 11 and Figure 8). Regarding the other three codes, objectivity prevails over subjectivity, but in differing degrees: mobility signs are mainly rendered objectively, editing signs are more often described subjectively, and finally the iconographic code shows a considerable number of subjective signs.

Therefore, in light of these results, we can state that objectivity and subjectivity do indeed coexist in AD scripts. Despite the objectivity recommendation encouraged by various guidelines, no purely denotational AD is to be found in our scripts. Instead, it seems that current filmic AD is the result of a combination of subjective and objective elements in different distributions. Moreover, this interrelation has been determined by applying multimodal analysis based on meaning codes as a tool that enables the operationalisation of objectivity, which confirms our hypothesis.

Precisely because of these results, one last remark is worth mentioning. As this paper hopes to prove, a key advantage of multimodal analysis is that it makes it possible to isolate very accurate pieces of information encapsulated in codes. In other words, multimodality allows the operationalisation of semiotic information. This operationalisation can then be used as a starting point for other sorts of research: cultural approximations (through which codes are cultural references usually transmitted?), contrastive approaches (is subjectivity more frequent in different languages?), reception studies (which subjective codes can users best accept?), etc.

All in all, multimodality has proved to be an approach with the potential to characterise the contents of AD scripts in terms of objectivity, which opens the doors to promising new research horizons.

## References

- 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]. <https://www.aenor.com/normas-y-libros/buscador-de-normas/UNE?c=N0032787>
- Baldry, A. (2000). *Multimodality and multimediality in the distance learning age*. Palladino Editore.
- Baldry, A., & Thibault, P. (2006). *Multimodal transcription and text analysis. A multimedia toolkit and coursebooks with associated on-line course*. Equinox.
- 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]. Tesis Doctorals en Xarxa (TXD) Repository. <https://www.tesisenred.net/handle/10803/669901#page=1>
- Betancor, D. (Executive Producer). (2018–2023). *Elite* [Élite] [TV series]. Netflix.
- Betancor, D. (Executive Producer). (2021). *Elite: Short Stories* [Élite: historias breves] [TV series]. Netflix.
- Braun, S. (2011). Creating coherence in audio description. *Meta: Translators' Journal*, 56(3), 645–662.
- Chaume, F. (2004). *Cine y Traducción* [Cinema and Translation]. Cátedra.
- Chaume, F. (2012). *Audiovisual translation: Dubbing*. Routledge.
- 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). Palgrave Macmillan. [https://doi.org/10.1007/978-3-030-42105-2\\_2](https://doi.org/10.1007/978-3-030-42105-2_2)
- 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>
- 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.
- Jakobson, R. (1959). *On linguistic aspects of translation*. In R. A. Brower (Ed.), *On translation* (pp. 232–239). Harvard University Press.
- Jankowska, A. (2015). *Translating audio description scripts. Translation as a new strategy of creating audio description*. Peter Lang.
- Kourdis, E. (2022). Semiotics. In F. Zanettin & C. Rundle (Eds.), *The Routledge handbook of translation and methodology* (pp. 139–154). Routledge. <https://doi.org/10.4324/9781315158945>
- Kress, G. (2014). What is mode? In C. Jewitt (Ed.) *The Routledge handbook of multimodal analysis* (pp. 60–75). Routledge.
- Kress, G., & Van Leeuwen, T. (1996) *Reading images: The grammar visual design*. Routledge.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.
- 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>

- 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). Routledge.
- Matamala, A. (2019). The VIW corpus: Multimodal corpus linguistics for audio description analysis. *RESLA. Revista Española de Lingüística Aplicada*, 32(2), 515–542.
- 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>
- Netflix. (2021). *Audio description style guide (v2.3)*. <https://partnerhelp.netflixstudios.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v2-3>
- Norris, S. (2004). *Analyzing multimodal Interaction: A Methodological framework*. Routledge.
- Ofcom. (2021). Ofcom’s code on television access services.  
[https://www.ofcom.org.uk/data/assets/pdf\\_file/0035/179954/tv-access-services-code.pdf](https://www.ofcom.org.uk/data/assets/pdf_file/0035/179954/tv-access-services-code.pdf)
- O’Halloran, K. (2012). Análisis del discurso multimodal [Multimodal discourse analysis]. *Revista Latinoamericana de Estudios del Discurso*, 12(1), 75–97.
- O’Halloran, K., Marissa K. L. E., & Tan, S. (2015). Multimodal semiosis and semiotics. In J. Webster (Ed.), *The Bloomsbury companion to M. A. K. Halliday* (pp. 386–411). Bloomsbury.
- Pina, Á. (Executive Producer). (2017–2021). *Money Heist [La casa de papel]* [TV series]. Netflix.
- Pina, Á. & Martínez Lobato, E. (Executive Producers). (2021–2023). *Sky Rojo [Sky Rojo]* [TV series]. Netflix.
- Remael, A., & Vercauteren, G. (2007). Audio describing the exposition phase of films. Teaching students what to choose. *TRANS, Revista de Traductología*, 11, 73–93.
- Reviere, N. (2018). Audio description in Dutch: A corpus-based study into the linguistic features of a new, multimodal text type [Unpublished doctoral dissertation]. University of Antwerp.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage Publications.
- 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.
- Stöckl, H. (2004). In between modes: Language and image in printed media. In E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality* (pp. 9–30). John Benjamins.
- 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>
- 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). Palgrave Macmillan. [https://doi.org/10.1007/978-3-030-42105-2\\_5](https://doi.org/10.1007/978-3-030-42105-2_5)
- Thibault, P., & Baldry, A. (2001). Towards multimodal corpora. In G. Aston & L. Burnard (Eds.), *Corpora in the description and teaching of English* (pp. 87–102). Clueb.
- Vercauteren, G. (2022). Narratology and/in audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 78–92). Routledge.
- Villanueva-Jordán, I. (2021). *La traducción de masculinidades gay en la teleficción: análisis multimodal del doblaje latinoamericano y peninsular de la serie Looking* [The translation of gay masculinities in telefiction: A Latin American and Spanish dubbing multimodal analysis of the series *Looking*] [Doctoral dissertation, Universitat Jaume I]. Tesis Doctorals en Xarxa (TXD) Repository. <https://www.tdx.cat/handle/10803/673040?locale-attribute=es>



- Walczak, A. (2017). *Immersion in audio description. The impact of style and vocal delivery on users' experience* [Doctoral dissertation, Universitat Autònoma de Barcelona]. Tesis Doctorals en Xarxa (TXD) Repository. <https://www.tdx.cat/handle/10803/402401#page=1>
- Zabalbeascoa, P. (2008). The nature of the audiovisual text and its parameters. In J. Diaz-Cintas (Ed.), *The didactics of audiovisual translation* (pp. 21–37). John Benjamins. <https://doi.org/10.1075/btl.77.05zab>