Journal of Audiovisual Translation

A Tiered Vocabulary System for Audio Description: Building Up Cinematic Jargon in Audio Introductions

José Luis Castillo-Flores ✉
Universidad de Las Palmas de Gran Canaria

Citation: Castillo-Flores, J. L. (2026). A Tiered Vocabulary System for Audio Description: Building Up Cinematic Jargon in Audio Introductions. *Journal of Audiovisual Translation*, *9*(1), 1–16. https://doi.org/10.47476/jat.v9i1.2026.370

Editor(s): E. Di Giovanni

Received: December 10, 2024

Accepted: August 11, 2025

Published: January 9, 2026

Abstract

In film audio description (AD), cinematic jargon is a fundamental component to portray the director's perspective and deliver an enriched cinematographic experience to the audience (Perego, 2014). Even though academia has partially tackled aspects related to how audio describers should approach film language (Remael et al., 2015), nothing has been said about how cinematic jargon should be arranged in the first place to offer professionals some structure to implement it effectively in their target text (TT).

In this study, we put forward a system that consists of a filmic analysis per sequence to locate technical cinematic terminology in Spanish, categorize it in tiers based upon the mainstays of vocabulary acquisition (Beck et al., 2013), and select eligible instances that fit in the cinematic audio-described script (CADS). The main objective is to establish a procedure that allows audio describers to incorporate purposeful cinematic terminology in their audio introduction (AI) and thus enhance blind and partially sighted (BPS) users' filmic experience. Our conclusion is that this system can contribute to conveying a unified glossary for cinematic AD purposes and help optimize the opportunity to disseminate this specialized knowledge in the long run.

Key words: audiovisual translation, accessibility, audio description, audio introduction, functionalism, cinematic language, tiered vocabulary system.

✉ luis.castillo@ulpgc.es, https://orcid.org/0009-0004-9726-9483

Introduction

Even though academia acknowledges the coexistence of multiple audio description (AD) styles that deviate from the standard practice (Szarkowska & Jankowska, 2024), one of the most contentious debates that has arisen in the landscape of film AD revolves around the use of cinematic language in the target text (TT) (Fryer & Freeman, 2013). In this regard, the conflict arises because most official guidelines recommend either avoiding or restricting the use of technical cinematic terminology (Perego, 2014, pp. 88–89). In contrast, a wave of scholars advocates for the use of cinematic jargon, given that "[i]n cinema, visual style (including elements such as camera work and editing techniques) is a major element of the film-watching experience" (Romero-Fresco & Fryer, 2013, p. 294). It is a matter that had not been brought up with the AD of live shows, since the sighted audience perceives "visual signs" (Zabalbeascoa, 2001) as a steady whole on stage and, consequently, has some freedom to address their eyesight as they please. Conversely, in cinema, visual signs are disclosed to the audience through a screen that serves as an intermediary, allowing for a fast, precise, and ever-changing construction of the film director's viewpoint. As Wilken and Kruger (2016) note:

> (…) film creates a story not simply by showing (as in theatre), but by showing from constantly shifting angles and perspectives, at varying distances from the object of the camera, and put together through editing in ways that suggest relations between successive shots. (p. 253)

Controversies aside, results from experiments conducted among blind and partially sighted (BPS) users have confirmed a generalized preference towards descriptions that include information related to cinematic language (Bardini, 2020b, 2022; Fryer & Freeman, 2013; Szarkowska, 2013; Walczak & Fryer, 2017). In sooth, what these studies underline is the need to explore feasible alternatives to conventional AD. In this direction, Bardini (2022) defines "cinematic AD" as follows:

> Interpretative AD, which strikes a balance between iconic description, the use of film terminology and interpretation of the film language. The use of technical terms primarily comes into play to describe specific cinematic elements, such as camera movements and editing techniques. Furthermore, when the describer deems it relevant, it also provides an interpretation of the meaning of the film techniques used. (p. 101)

This AD style poses manifold challenges for the audio describer, the most fundamental being how to generate an effective cinematic audio-described script (CADS) that will necessarily require an explanation of technical terminology, given that time restrictions for oral inserts are, by nature, very tight. A plausible solution could be through an audio introduction (AI), an added oral text that is played before the beginning of a movie to provide the BPS audience with extra information about the piece, including concepts related to its cinematic nature. Heretofore, many different aspects of AIs for film AD have been discussed, including audience reception (Di Giovanni, 2014; Romero-Fresco & Fryer, 2013; Szarkowska & Jankowska, 2015), recommended structure (Fryer & Romero-Fresco, 2014), correlation with the audio-described script (Fryer, 2016), or text functions (Romero-Fresco, 2022). Nonetheless, little has been said about how audio describers should proceed when incorporating specialized cinematic jargon into their AIs, which is a key aspect in guiding professionals

through the AD process and ultimately leveraging this tool to enhance users' enjoyment and appreciation of the filmmaker's style.

In this paper, we suggest a tiered vocabulary system that serves as a starting point to organize technical cinematic terminology and help audio describers discern what concepts are eligible to be implemented in their AI. If the final objective is to optimize the space available in AIs to foster specialized knowledge among BPS users and ease TT comprehension, it seems appropriate to find shelter under the umbrella of functionalism, which comprises the next section of this article.

1. Functionalism as a Framework for Cinematic Audio Description

Stemming from the work of scholars such as Reiss (1977/1989), Vermeer (1978/1989) and Nord (2013), functionalism posits that translation choices should be guided by the communicative intent and needs of the target audience. The Skopos theory, which underlines "purpose" as the key principle during the decision-making process in translation, is central to this framework. As Vermeer (1978/1989) states:

> Each text is produced for a given purpose and should serve this purpose. The skopos rules thus reads as follows: translate/interpret/speak/write in a way that enables your text/translation to function in the situation in which it is used and with the people who want to use it and precisely in the way they want it to function. (as translated by Nord, 2013, p. 204)

This approach was adopted during the European project ADLAB (2011–2014), which resulted in the publication of a complete set of research-based AD guidelines (Remael et al., 2015). In the introduction, the editors move away from the notion of "equivalence" between the source text (ST) and TT, given that it is a complex matter to define, and never absolute. As they note: "In the case of AD the concept is even more problematic: when is the verbal rendering of an audio-visual product 'equivalent' to its aural/visual ST?" (Remael et al., 2015, p. 15). A year before, Vercauteren (2014), who was involved in the creation of these guidelines as editor and contributor, suggested systematizing research in AD by integrating it under the functionalist paradigm, away from the traditional concept of "equivalence". By that time, Nord (2013) had already clarified that "Equivalence is a static, result-oriented concept describing a relationship of 'equal communicative value' between two texts or, on lower ranks, between words, phrases, sentences, syntactic structures and so on." (p. 204). Considering the multimodal, intersemiotic nature of AD, and more specifically its infrasemiotic dimension (Gottlieb, 2017), in which information available through two channels, the visual and the auditory, must be conveyed solely through audio, finding a direct equivalence indeed seems unfeasible. Instead, Nord (2013) advocates for the notion of "adequacy" (p. 204), which entails that the TT must be adapted to the requirements of the "translation brief".

Departing from this theoretical foundation, a growing number of scholars have delved into functionalism to frame their research in AD over the past decade. Concomitant to the proposal of Vercauteren mentioned above, Mazur (2014) states that in accordance with the skopos theory "The

choice of the global strategy and the subsequent procedures and (local) strategies will depend on a number of factors, such as the text type, the target audience, or the purpose and function of translation (…)" (p. 186). In the context of AD, this functionalist perspective allows for greater flexibility and creativity in rendering visual content into verbal form. Three years later, Bardini (2017) further reinforces this impression for alternative AD styles: "Functional theory provides a solid framework for developing new audio description styles that deviate from conventional guidelines but take users' needs and source texts' properties into account" (p. 67). In 2020, the same author (Bardini, 2020a) creates a taxonomy to classify "AD techniques" (ADT) based on the "dynamic and functionalist approach" previously developed by Molina and Hurtado (2002). Bardini (2020a) defines ADT as "options available to audio describers to render the audiovisual message into the verbal system when audio describing a fragment of an audiovisual text" (p. 278), and describes them as functional, combinable and determining towards the AD script at a micro-level. That same year, Mazur (2020) suggests a three-layered analysis (contextual, macrotextual, and microtextual) of the ST to determine applicable macro and micro strategies for the AD. In this same article, inspired by the functional classification established by Reiss (1977/1989), Mazur also establishes the five functions that shape multimodal texts: informative, narrative, expressive, persuasive and entertaining. Two years later, Marra (2022) applies a functionalist approach to the analysis of the Italian and Spanish ADs of different scenes from the same film. The goal is to prove how specific linguistic choices can alter the result of an AD and, consequently, the communicative aim of the TT. Lastly, Mazur (2023) applies a functional model of AD to determine macro and micro strategies with grounds in the functional priorities of audiovisual content. The proposal adapts the principles of the skopos theory for AD in the form of four rules that follow a hierarchal order: (1) the skopos rule, which safeguards the purpose of the AD; (2) the cohesion rule, which watches over cohesion in the global audio-described text; (3) the coherence rule, which guarantees a coherent TT for the BPS audience; and (4) the correspondence rule, which looks after intersemiotic correspondence between ST and TT. The article also illustrates how more than one type of AD can be generated according to the target audience and the purpose of the TT.

Turning back to the content included in the ADLAB guidelines, it should be noted that there is a section on film language (Remael et al., 2015, pp. 30–35) where concrete indications are given to identify which categories for film techniques exist (mise-en-scène, cinematography and editing), what their functions are within the audiovisual text (denotative, expressive, symbolic or aesthetic), and how they can be reflected in the AD (only naming the technique, naming the technique and describing its function or directly describing its function or meaning)[1]. On this same publication, there is also a segment dedicated to AIs (Remael et al., 2015, pp. 58–61) that covers some previous considerations (if the AI is going to be run independently or combined with an AD script, if the AD script was previously written, and if it is going to be recorded or delivered live), its functions (informative, foreshadowing, expressive-aesthetic, instructive), and how to craft an AI (structure,

[1] This section was written by Elisa Perego, from the Università di Trieste.

content and style)[2]. However, these guidelines were conceived for general use and do not contemplate alternative AD styles, so it does not focus on how to build cinematic jargon in the AI.

All the approaches described above share the same functionalist grounds. In our case, our goal is to create a cinematic AD that renders the visuals so that the target audience can embrace the cinematic qualities of a film and, to reach this purpose, including an AI becomes essential. In fact, this statement by Fryer (2016) reaches new heights: "In some cases, the AI can also be used as a kind of glossary, explaining technical terms in advance, allowing the strategy of naming without explicitation to be used in the Dynamic AD" (p. 159). From our perspective, AIs constitute a critical tool to define technical terminology that will eventually be put into practice in the dynamic CADS. Without this component, there would not be any available space to explain key concepts that are going to be named later, and successful communication would not be accomplished. If the skopos is to deliver a comprehensible cinematic AD for BPS users, and the only way to accomplish such a goal effectively is by using an AI equipped with definitions for specialized jargon, then a suitable categorization system for technical cinematic terminology is needed.

2. A Tiered Vocabulary System for Audio Description

Originally designed for language acquisition purposes, the tiered vocabulary system is a model to categorize words in accordance with their meaning, specificity and complexity developed by Beck et al. (2013). In education, a tiered vocabulary system is just a basic step followed by teachers who apply scaffolding techniques as part of their methodology. Scaffolding is all about delivering lessons in segments, progressively providing less support while students master a particular material and grow their confidence. While we acknowledge that the teacher-student relationship does not apply to that of an audio describer and a user, we believe that a similar scaffolding strategy can enhance AIs for cinematic AD: it can name and define technical terms, help BPS users become familiar with cinematic AD, and eventually enable them to enjoy and appreciate the film director's style.

In the system proposed by Beck et al. (2013) categorization is divided into three tiers. Tier 1 is for words that the user is already familiar with. Tier 2 contains high-utility words that will become the main target of instruction. Tier 3 refers to domain-specific words that make sense when learning a specialized subject matter. One of the main differences between other previous works on vocabulary selection (Carroll et al., 1971; Zeno et al., 1995; Hiebert & Kamil, 2005) and the model proposed by Beck et al. (2013) is that the latter does not focus on the frequency of words solely:

> Frequency lists can seem very tempting, because they are objective − gathered from actual data on use of words in the language. But this very fact is also why they should not be relied on. Frequency merely indicates how often a word appears in print compared to other words

[2] This section was written by Nina Reviers, from the University of Antwerp.

> in the language. That fact does not exactly translate into how difficult a word is or even how useful it is to a user's repertoire. (p. 22)

From their perspective, criteria for identifying potential Tier 2 words are based on "importance and utility", which encompasses general frequency; "conceptual understanding", which covers learners' ability to grasp a concept and then describe it with a greater level of precision and specificity, and "instructional potential", which comprehends its practical usefulness in context (Beck et al., 2013, p. 28). These criteria provide a useful lens for considering how cinematic terminology can be purposefully introduced to BPS audiences, rather than relying only on frequency or familiarity. However, unlike a classroom setting, AD only occurs once, with no opportunity for interaction, clarification or repetition. Therefore, users cannot ask questions, receive feedback or pause to reflect. This is why the integration of specialized cinematic jargon into a contextualized, time-constrained communicative act poses a daunting challenge. To address this, an AI can operate as a scaffolding mechanism so that users do not encounter unfamiliar terms during the film. This approach helps bridge the gap between specialized cinematic jargon and user comprehension, while respecting the natural constraints of AD.

Bearing in mind that cinematic language is central to conveying the visual style, and that our goal is to provide an accessible, enjoyable and comprehensible filmic experience for a diverse audience, we propose rethinking the original tiered model and drawing the Tier 3 category to serve as the standard pool of vocabulary words we depart from. Accordingly, our categorization will be established as follows:

- **Tier 1:** words that can be understood implicitly by the semantic relationship established between the name itself and its actual meaning. These terms will not be developed in the AI.
- **Tier 2:** words that represent our main target and require an explicit instructive definition in our AI.
- **Tier 3:** words that can be considered advanced technical vocabulary. They could be added to the AI at the audio describer's discretion according to its aesthetic relevance within the film and depending on the overall cognitive load imposed on our BPS audience.

From these ideas on how to arrange terminology for cinematic AD, we have created a tiered vocabulary system that is divided into three stages: location, categorization, and selection. We would like to note that the model presented below will focus exclusively on technical cinematic terminology related to camera position and movement and editing techniques. While we are aware that other valuable constituents portray a film director's viewpoint, such as mise-en-scène (costume, lighting or colour, to mention just a few), we are particularly interested in fixed cinematic terms that could be common to any film. To some extent, our model assumes that audio describers possess a basic understanding of cinematography, such as camera angles and movements, editing techniques, and other elements of film language to identify and incorporate relevant terminology into the AD script. However, we are aware that not all practitioners may have in-depth knowledge in this area. For this

reason, we strongly recommend referring to additional resources such as specialized glossaries, relevant publications on the film director's style, or collaboration with film experts. Finally, the model is designed to be adaptable: it can support both experienced AD professionals and those still building their cinematic literacy.

To illustrate our model, we have drawn examples from three sequences of *In the Mood for Love* (2000), by Chinese filmmaker Wong Kar-wai. The cinematic AD was self-crafted in Spanish, so vocabulary terms will be provided in this language, along with a tentative back translation into English that has been added only the first time a term is mentioned. Additional examples used to illustrate other issues that are not reflected in the excerpts are provided in both Spanish and English.

2.1. Stage 1: Location

On this first leg of the process, audio describers should thoroughly observe the ST on the lookout for two key aspects of cinematic language, namely the camera position and movement, and the editing techniques. As they perceive these, they should jot down the corresponding technical cinematic terminology that renders the visuals and the number of potential occurrences for each concept in the CADS. This vocabulary list will set the grounds to categorize terms in tiers afterwards.

Table 1

Location of Cinematic Terms in Spanish

Excerpts	Vocabulary List	Potential occurrences
Sequence #1	Travelling (tracking shot)	5
00:00:00-	Primer plano (close-up shot)	8
00:05:57	Plano de espaldas (back-to-the-camera shot)	1
	Fundido en negro (fade to black)	2
	Plano medio (medium shot)	9
	Plano general (long shot)	2
	Contrapicado (low-angle shot)	1
	Plano de detalle (big close-up shot)	1
	Toma sin cortes (take without cuts)	1
	Fuera de campo (out of frame)	1
	Travelling	4
	Plano de espaldas	2
	Encuadre vacío (empty shot[3])	1

[3] In Anglophone film studies, the term "pillow shot" is widely accepted to name this technique. However, we have opted for the use of "empty shot" in consideration to its original sense that stems from Chinese philosophy and the fact that the film chosen for our analysis is also Chinese. For a richer discussion on the topic, see Duan (2021).

Excerpts	Vocabulary List	Potential occurrences
Sequence #2	Primer plano	21
00:24:32-	Plano medio	10
00:31:15	Plano de detalle	6
	Contrapicado	1
	Plano de situación (establishing shot)	1
	Plano general	2
	Barrido rápido (fast pan)	3
Sequence #3	Plano general	3
01:06:07-	Plano medio	4
01:13:52	Travelling	2
	Primer plano	14
	Plano de detalle	2
	Fundido en negro	2
	Corte en J (J-cut)	1
	Plano de espaldas	1

Source. Author's own work.

In these three sequences, a total of 14 different specialized concepts have been registered. At this point, naming simplification becomes essential. When consulting literature about cinema, we often encounter terms that are used interchangeably to designate the exact same effect, which could be confusing for any apprentice of a specialized subject. Some examples in Spanish are "plano aéreo" vs. "plano a vista de pájaro" or "plano holandés" vs. "plano aberrante". In English, some instances are "dolly shot" vs. "tracking shot" or "sequence shot" vs. "long take" vs. "one-shot take". Although each of them individually could be easily understood by a BPS audience, it is advisable to strive for a unified jargon for cinematic AD purposes. Some cases seem to be more delicate, as the example of "empty shot" vs. "pillow shot" points out.

2.2. Stage 2: Categorization

In this next layover, vocabulary words should be arranged in tiers, and the number of potential occurrences in the dynamic CADS should be reflected in parentheses. As we mentioned earlier, the Tier 1 category will include terms that do not require an explanation in our AI because their meaning can be understood implicitly by the name alone. The Tier 2 category will be comprised of terms that require an explicit explanation. These words represent our main target and could potentially be included in our AI. The Tier 3 category will host more complex, advanced technical terms that will also call for an explicit elaboration, but that must be carefully evaluated by the practitioner based on its aesthetic relevance and the overall cognitive load imposed on the BPS audience.

At this point, it is important to note that the classification of terms into tiers may not map consistently across languages due to semantic and lexical differences. This means that vocabulary categorization may differ from one language to another. For instance, in English, the concept "travelling shot" could be considered a Tier 1 term because it can be understood implicitly without further elaboration, whereas in Spanish, the term calls for a full explanation in the AI. Subsequently, our tiered vocabulary system will need to be adapted to each language and, therefore, involve collaboration with AD professionals and linguists from different countries to refine tier categorizations based on linguistic and cultural aspects.

Table 2

Categorization of Cinematic Terms in Tiers

Tier 1	Tier 2	Tier 3
Primer plano (33)	Travelling (11)	Corte en J (1)
Plano de espaldas (4)	Fundido en negro (4)	
Plano medio (23)	Contrapicado (2)	
Plano general (7)	Fuera de campo (1)	
Plano de detalle (9)	Encuadre vacío (1)	
Toma sin cortes (1)	Barrido rápido (3)	
Plano de situación (1)		

Source. Author's own work.

As the table shows, there are seven concepts that fall in the Tier 1 column: "primer plano", "plano de espaldas", "plano medio", "plano general", "plano de detalle", "toma sin cortes", and "plano de situación". If time restrictions allow it, they can be named directly in the CADS. In general lines, audio describers should be practical and avoid unnecessary repetitions. For instance, if in a scene shot in consecutive close-ups, two characters are holding a conversation and there is a constant shift of angle, it will not be necessary to indicate the type of shot every single time. There are six concepts in the Tier 2 column: "travelling", "fundido en negro", "contrapicado", "fuera de campo, "encuadre vacío", and "barrido rápido". Each of these could potentially be defined in our AI. Finally, there is only one concept that we have categorized in the Tier 3 column. In this case, "corte en J" is an editing technique in which the audio of the next sequence precedes the image change, a notion that could be considered hard to perceive in a medium in which timing is often altered to fit oral inserts.

2.3. Stage 3: Selection

In this part of the process audio describers should determine which vocabulary words from Tier 2 – and Tier 3 exceptionally –, will be included in the AI. Inspired by the model designed by Beck et al. (2013, p. 37), audio describers will need to take into consideration three mandatory conditions: (1) the concept represented is understandable; (2) the concept can be explained in simple language,

and receivers could potentially be able to rephrase its meaning with their own words; (3) the concept is going to be put into practice shortly after in the dynamic CADS. This last requirement is significant because every time the term appears in the dynamic CADS, it represents an opportunity for the BPS audience to recognize and appreciate the film director's stylistic choices, thereby deepening their engagement and enjoyment in relation to the cinematic aspects. When these conditions are clear, audio describers need to work along with their drafted CADS to identify the available space in oral inserts and include technical cinematic terminology in their dynamic AD. If necessary, they can consider compressing information or disregarding less-relevant factual aspects to adapt their TT to the translation brief. In this sense, they should seek a balance and embed the specialized jargon related to camera shots and movements and editing effects in a natural way in the CADS. To keep track of the opportunities for vocabulary recognition given to the audience, we have included an account of the effective occurrences that will be present in the script.

Table 3

Selection of Cinematic Terms That Fit in the CADS

Tier 2	Effective Occurrences	Tier 3	Effective Occurrences
Travelling (11)	11	Corte en J (1)	0
Fundido en negro (4)	3		
Fuera de campo (1)	1		
Encuadre vacío (1)	1		
Contrapicado (2)	0		
Barrido rápido (3)	2		

Source. Author's own elaboration.

In this case, five concepts will be defined in our AI and two will be excluded. The Tier 2 term "contrapicado" has only two occurrences, but none of them fit in the CADS. The Tier 3 term "corte en J" appears in one occasion and will be left out on three counts: first, its meaning is difficult to grasp given the context; second, it is combined with a fade to black effect and the verbal representation of two or more simultaneous strategies could impose a strong cognitive load on an audience that is unfamiliar with cinematic terminology; and third, there is not enough time to express one of the effects in the CADS. Apart from that, the BPS audience will also lose one opportunity for vocabulary recognition for the terms "fundido en negro" and "barrido rápido" that account for a total of 3 and 2 effective occurrences, respectively. Regarding the total number of vocabulary words, we believe that audio describers should include every Tier 2 term in their AIs that meets the three conditions detailed above. In consonance with Fryer & Romero-Fresco (2014): "Unlike most AD scripts, AIs do not need to shy away from using specialised or technical terms such as camera angles or editing techniques, as these can be mentioned with the resulting visual effect explained (…)" (p. 24). Even though we are aware that a cinematic AD entails a greater cognitive effort for a BPS

audience, particularly when it includes unfamiliar technical terms, we agree with Perego (2014) when she says that

> In the first place, the range of technical terms to be included in a film, and the average range of editing techniques used, is typically limited; in the second place, exposure to such terms is the first step to noticing, learning and eventually enjoying them. (p. 96)

However, we recognize that exposure alone may not be enough to support the integration of these concepts, especially given the wide range of preferences, levels of cinematic knowledge, and cognitive processing abilities among the BPS audience. This is why we propose using an audio introduction as a tool to scaffold key cinematic terms, since it allows users to engage with potentially unfamiliar terminology in a context that is not affected by other elements. In this sense, our approach, rather than assuming a uniform response, acknowledges the diversity of BPS users while still advocating for inclusive access to the stylistic aspects of film.

2.4. Building Up Cinematic Jargon in Audio Introductions

Once the selection process is completed, audio describers can proceed to define the terms in the AI. In general, language should run smoothly, syntax should be kept short and simple with a preference for the present tense and coordinated sentences, as Fryer & Romero-Fresco (2014) have remarked. Note that the fragment below corresponds exclusively to aspects related to camera position and movement and editing techniques, as we clarified at the end of section 2.

> En varios momentos del filme se emplea la técnica del **travelling**, que consiste en un desplazamiento lateral de la cámara y que, normalmente, acompaña a un personaje mientras camina. Puntualmente, se recurre al **fundido en negro**, que es una transición en la que tanto la imagen como el sonido se atenúan, hasta que la pantalla se queda totalmente en negro. En ocasiones se produce un **encuadre vacío**, en donde la cámara mantiene el enfoque fijo en un espacio estático. En ciertas instancias, los personajes que participan en la secuencia están **fuera de campo**, es decir, presentes en el espacio de la escena, pero no visibles en pantalla. En unos puntos determinados de la película se emplean unos **barridos rápidos**, es decir, unos arrastres bruscos de cámara. [Original in Spanish]

> In several moments of the film, the **tracking shot** technique is used, which consists of a lateral movement of the camera that normally accompanies a character who is walking. Occasionally, a **fade to black** is applied, which is a transition in which both the image and the sound progressively disappear until the screen goes completely black. Sometimes an **empty shot** occurs, where the camera maintains its focus fixed on a static space. In certain instances, the characters participating in the sequence are **out of frame**, which means that they are present in the space of the scene, but not visible on screen. At certain points in the film, sudden camera drags called **fast pans** are used. [Back translation into English]

3. Conclusion

In this article, we have set forth a tiered vocabulary system to pin down cinematic jargon within a film, categorize it in tiers to help audio describers discern what concepts are eligible to be implemented in their AI, and choose eligible instances to optimize the space available in the AI with useful terminology that will be put into practice in the dynamic CADS. While we believe our model offers a useful framework for incorporating cinematic terminology into AD practice, we also recognize its limitations. The success of this approach depends primarily on the diversity of the BPS audience, whose prior exposure to film language, cognitive processing abilities, and preferences may vary significantly. In addition, time constraints can also limit the presence of cinematic terminology in the dynamic AD. We also acknowledge that the model presented here is open for growth. For instance, making the vocabulary model applicable across international AD contexts requires a more systematic consideration of cross-linguistic variability. Apart from this, it could be enhanced by adding specific vocabulary related to other key constituents of film language, such as elements of mise-en-scène (costume, lighting, colour, etc.). In the short term, we think that our system can ease audio describers' decision-making process and gradually aid in the creation of a common unified glossary for cinematic AD purposes. This latter goal could help reduce the audio describer's workload and stimulate BPS users' interest in cinematography and appreciation for cinematic language over time. We are fully aware that a tiered vocabulary system is just a stage in the greater scaffolding strategy of a learning experience, and that watching a film is not the same as being in a classroom. However, an audio introduction plays a central role in improving BPS users' knowledge, engagement and enjoyment of cinematic aspects. After all, the skopos of a cinematic AD is to deliver a comprehensible TT that renders the visuals, and the only way to achieve that is by providing extra support. What is more, there are multiple gains towards promoting cinematic AD among BPS users: they can learn new concepts and find new motivations, get a deeper and more faithful understanding of the filmmaker's viewpoint, improve their memory skills, boost their autonomy by applying concepts to the dynamic CADS, develop their imagination in ways they might not have experienced before, and bridge gaps amid patrons with different sight loss conditions.

Be that as it may, we view this model as a flexible starting point that can evolve in dialogue with AD users, practitioners, and scholars alike. Nonetheless, many questions that are crucial for the future of cinematic AD remain unanswered: Are BPS users willing to know about the specifics of the film director's viewpoint as they watch the movie? To what extent are they eager to learn about cinematic language? How far is enough before the CADS becomes tiring? Will BPS users voluntarily play an AI before a movie? How can we make them aware that AIs are a fundamental tool to develop a CADS? How can we make AIs more appealing to BPS users? Can cinematic knowledge be complemented beyond the reach of AD? These questions point to the need for further empirical research involving BPS audiences to evaluate the effectiveness of our model in practice. One promising path would be to conduct experimental research applying our tiered vocabulary system to a selected corpus and the same BPS subjects to assess shifts in preferences, levels of cinematic knowledge, and perceived cognitive load over time. In this sense, it would be interesting to include alternatives that go beyond

the reach of audio describers as well, such as working with organizations for the blind to promote workshops on cinematic language. Other prospect studies could focus on adapting our model across different languages and cultural contexts. Another potential strand could explore the emotional and perceptual differences between how sighted users interpret specific camera angles or editing techniques in contrast with how these same components are received by BPS audiences through their verbal AD.

References

Bardini, F. (2017). Audio description style and the film experience of blind spectators: Design of a reception study. *Rivista internazionale di tecnica della traduzione / International journal of translation*, *19*, 49–73. https://doi.org/10.13137/2421-6763/17351

Bardini, F. (2020a). Audio description and the translation of film language into words. *Ilha do Desterro*, *73*(1), 273–295. https://doi.org/10.5007/2175-8026.2020v73n1p273

Bardini, F. (2020b). Film language, film emotions and the experience of blind and partially sighted viewers: a reception study. *The Journal of Specialised Translation*, *33*, 259–280. https://doi.org/10.26034/cm.jostrans.2020.556

Bardini, F. (2022). Audiodescripción, subjetividad y experiencia fílmica. *Quaderns del CAC*, *48*, 97–106. https://doi.org/10.34810/qcac48id406116

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

Carroll, J., Davies, P., & Richman, B. (1971). *Word frequency book*. American Heritage.

Di Giovanni, E. (2014). Audio introduction meets audio description: An Italian experiment. In R. M. Bollettieri Bosinelli, E. Di Giovanni, & L. Rossato (Eds.) Across screens across boundaries [Special issue]. *InTRAlinea*, https://www.intralinea.org/specials/article/2072

Duan, S. (2021). Thinking, feeling and experiencing the 'empty shot' in cinema. *Film-Philosophy, 25*(3), 346–361. https://doi.org/10.3366/film.2021.0179

Fryer, L. (2016). Audio introductions. In L. Fryer (Ed.), *An introduction to audio description: A practical guide* (pp. 155–163). Routledge. https://doi.org/10.4324/9781315707228

Fryer, L., & Freeman, J. (2013). Cinematic language and the description of film: Keeping AD users in the frame. *Perspectives: Studies in Translatology*, *21*(3), 412–426. https://doi.org/10.1080/0907676X.2012.693108

Fryer, L., & Romero-Fresco, P. (2014). Audio introductions. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio description: New perspectives illustrated* (pp. 11–28). John Benjamins. https://doi.org/10.1075/btl.112.02fry

Gottlieb, H. (2017). Semiotics and translation. In K. Malmkjaer (Ed.), *The Routledge handbook of translation studies and linguistics* (pp. 45–63). Routledge. https://doi.org/10.4324/9781315692845

Hiebert, E., & Kamil, M. (2005). *Teaching and learning vocabulary: Bringing research to practice*. Lawrence Erlbaum.

Marra, L. (2022). Audio description (AD) and multiple levels of meaning in films: A functionalist perspective. *Lingue e linguaggi*, *53*, 303–319. https://doi.org/10.1285/i22390359v53p303

Mazur, I. (2014). Gestures and facial expressions in audio description. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio description: New perspectives illustrated* (pp. 179–197). John Benjamins. https://doi.org/10.1075/btl.112.11maz

Mazur, I. (2020). A functional approach to audio description. *Journal of Audiovisual Translation*, *3*(2), 226–245. https://doi.org/10.47476/jat.v3i2.2020.139

Mazur, I. (2023). Same film, different audio descriptions: Describing foreign films from a functional perspective. *Universal Access in the Information Society*, *23*, 511–521 https://doi.org/10.1007/s10209-023-00998-9

Molina, L., & Hurtado Albir, A. (2002). Translation techniques revisited: A dynamic and functionalist approach. *Meta*, *47*(4), 498–512. https://doi.org/10.7202/008033ar

Nord, C. (2013). Functionalism in translation studies. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 201–212). Routledge. https://doi.org/10.4324/9780203102893

Perego, E. (2014). Film language and tools. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio description: New perspectives illustrated* (pp. 81–101). John Benjamins. https://doi.org/10.1075/btl.112

Reiss, K. (1989). Text types, translation types and translation assessment (A. Chesterman, Trans.). In A. Chesterman (Ed.), *Readings in translation theory* (pp. 105–115). Finn Lectura. (Original work published 1977).

Remael, A., Reviers, N., & Vercauteren, G. (Eds.). (2015). *Pictures painted in words: ADLAB audio description guidelines*. *Edizioni Università di Trieste*. https://www.openstarts.units.it/entities/publication/be626243-8361-4b94-a74e-05224e6120ae/details

Romero-Fresco, P. (2022). Audio introductions. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 423–433). Routledge. https://doi.org/10.4324/9781003003052-33

Romero-Fresco, P., & Fryer, L. (2013). Could audio-described films benefit from audio introductions? An audience response study. *Journal of Visual Impairment & Blindness*, *107*(4), 287–295. https://doi.org/10.1177/0145482X1310700405

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/10.1177/0145482X1310700

Szarkowska, A., & Jankowska, A. (2015). Audio describing foreign films. *The Journal of Specialised Translation*, *23*, 243–269. https://doi.org/10.26034/cm.jostrans.2015.348

Szarkowska, A., & Jankowska, A. (2024). *Introducing audiovisual translation*. Routledge. https://doi.org/10.4324/9781003038122

Vercauteren, G. (2014). A translational and narratological approach to audio describing narrative characters. *Traduction, textes, médias*, *27*(2), 71–90. https://doi.org/10.7202/1037746ar

Vermeer, H. (1989). Skopos and commission in translational action (A. Chesterman, Trans.). In A. Chesterman (Ed.), *Readings in translation theory* (pp. 173–187). Finn Lectura. (Original work published 1978). Reprinted from *The translation studies reader*, by L. Venuti, Ed., 2000, Routledge.

Walczak, A., & Fryer, L. (2017). Creative description: The impact of audio description style on presence in visually impaired audiences. *British Journal of Visual Impairment*, *35*(1), 6–17. https://doi.org/10.1177/0264619616661603

Wilken, N., & Kruger, J.-L. (2016). Putting the audience in the picture: Mise-en-shot and psychological immersion in audio described film. *Across Languages and Cultures*, *17*(2), 251–270. https://doi.org/10.1556/084.2016.17.2.6

Zabalbeascoa, P. (2001). El texto audiovisual: factores semióticos y traducción [The Audiovisual Text: Semiotic Factors and Translation]. In *¡Doble o Nada!: Actas de las I y II jornadas de doblaje y subtitulación* [*Double or Nothing: Proceedings of the First and Second Seminars on Dubbing and Subtitling*] (pp. 113–126). Universitat Jaume I.

Zeno, S., Ivens, S., Millard, R., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science.